Today: Outline

- **Logistics**
- **Model Selection and Evaluation**
- **Explainable AI**

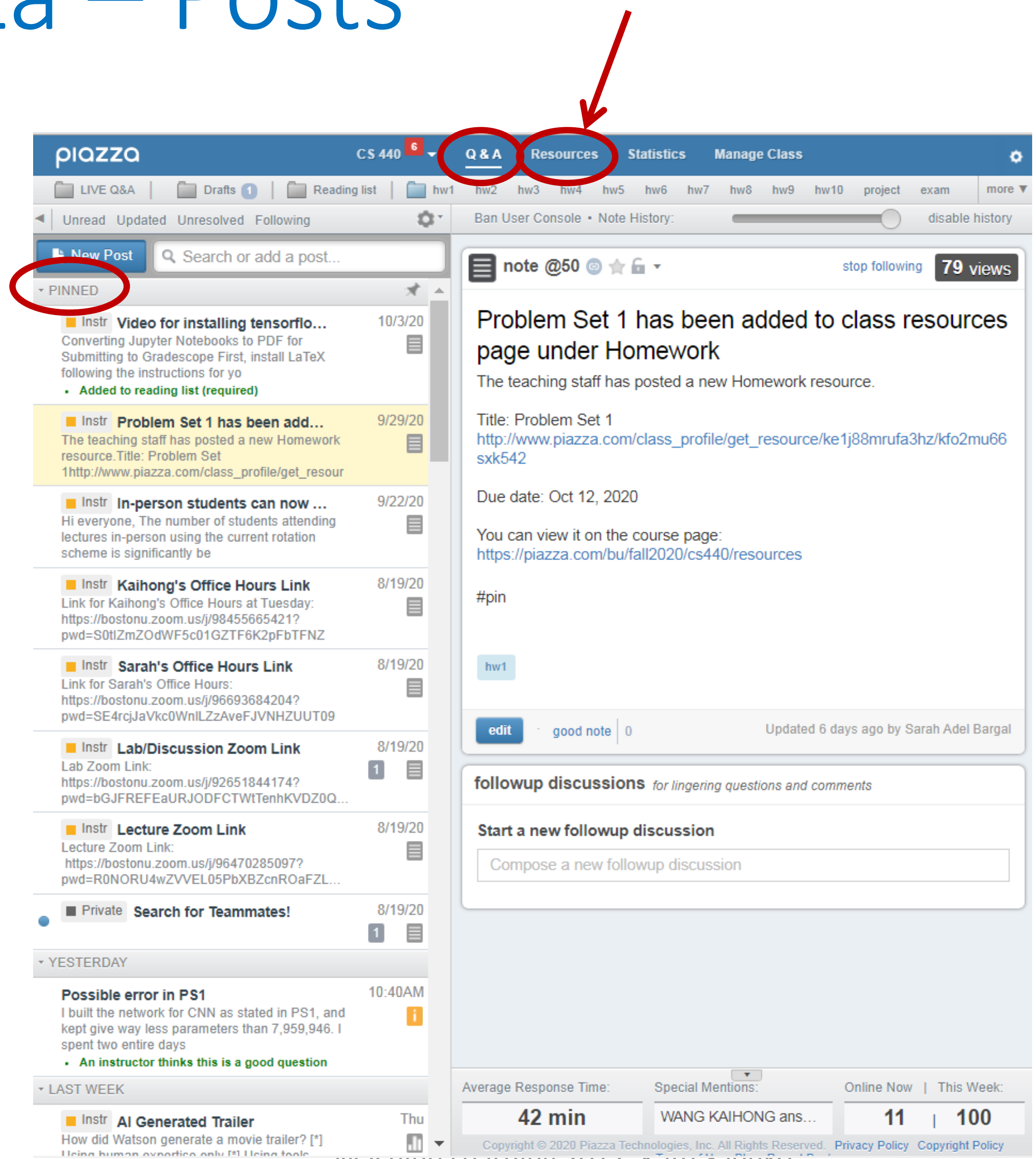
- Reminders:

  - Problem Set 1, due: Oct 12 by midnight
  - Midterm Exam, in class, Oct 20 (Practice problems will be posted)

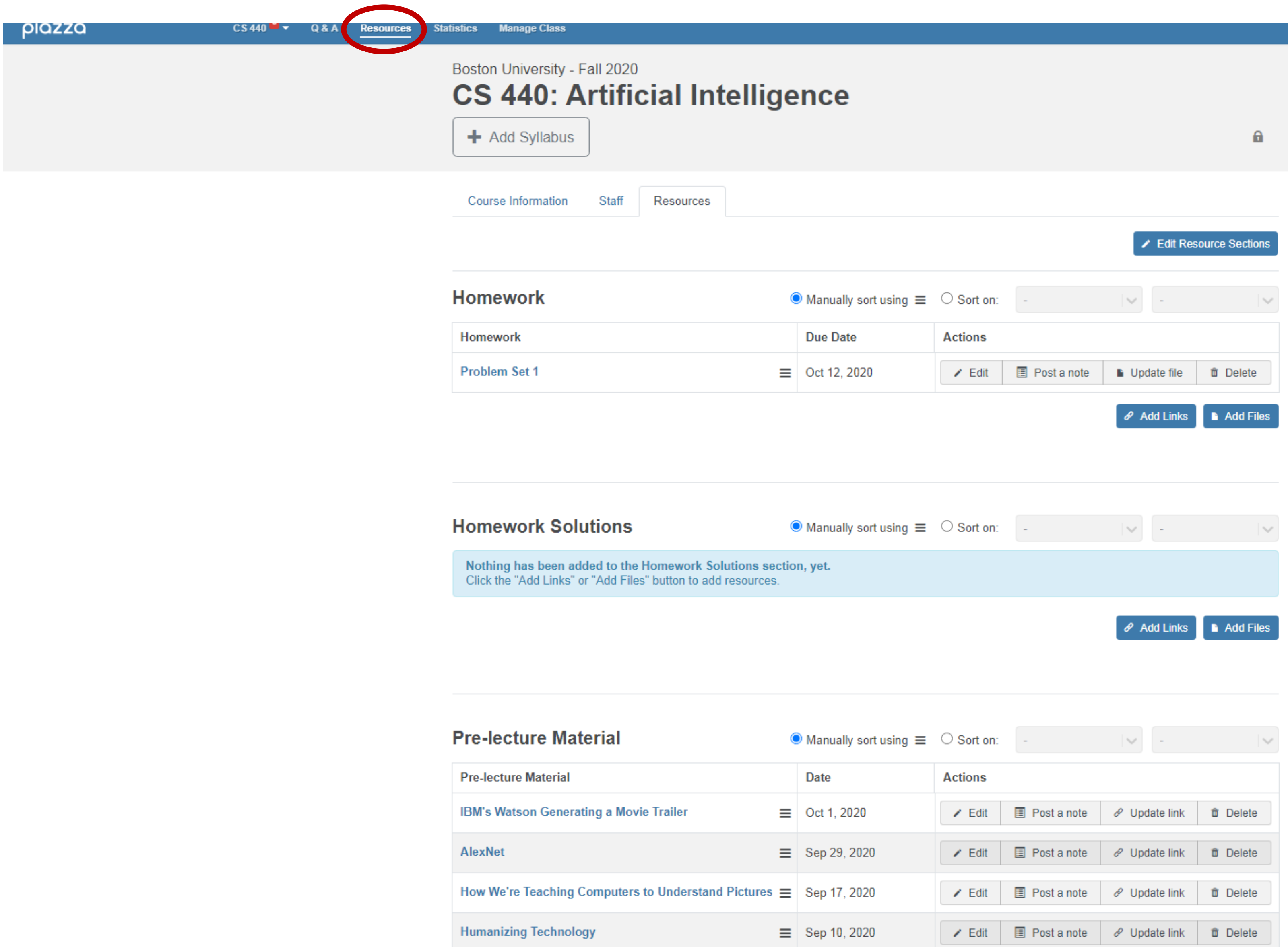
- Announcements:

  - Lab on Fri Oct 9 will be a chance to get help with PS1
  - No class on Oct 13 per BU Calendar (Substitute Mon Schedule of Classes)

Piazza – Posts



Piazza - Resources



Lecture Notes

Manually sort using ≡ Sort on: - -

Lecture Notes	Lecture Date	Actions
Lecture 9: Video, Password: ICQ!.X2c	Oct 1, 2020	Edit / Post a note / Update link / Delete
Lecture 9: Neural Networks V	Oct 1, 2020	Edit / Post a note / Update file / Delete
Lecture 8: Video, Password: x!8x+V?$	Sep 29, 2020	Edit / Post a note / Update link / Delete
Lecture 8: Neural Networks IV	Sep 29, 2020	Edit / Post a note / Update file / Delete
Lecture 7: Video, Password: RhT^+IX0	Sep 24, 2020	Edit / Post a note / Update link / Delete
Lecture 7: Neural Networks III	Sep 24, 2020	Edit / Post a note / Update file / Delete
Lecture 6: Video, Password: 7!VG.88g	Sep 22, 2020	Edit / Post a note / Update link / Delete
Lecture 6: Neural Networks II	Sep 22, 2020	Edit / Post a note / Update file / Delete
Lecture 5: Video, Password: DyA&1MTF	Sep 17, 2020	Edit / Post a note / Update link / Delete
Lecture 5: Neural Networks I	Sep 17, 2020	Edit / Post a note / Update file / Delete
Lecture 4: Video, Password: N7fSvw8!	Sep 15, 2020	Edit / Post a note / Update link / Delete
Lecture 4: Learning from Examples II	Sep 15, 2020	Edit / Post a note / Update file / Delete
Lecture 3: Video, Password: ctq$1ShM	Sep 10, 2020	Edit / Post a note / Update link / Delete
Lecture 3: Learning from Examples I	Sep 10, 2020	Edit / Post a note / Update file / Delete
Lecture 2: Video, Password: 44Wq9Z8&	Sep 8, 2020	Edit / Post a note / Update link / Delete
Lecture 2: Ethics in AI	Sep 8, 2020	Edit / Post a note / Update file / Delete
Lecture 1: Video, Password: j=5HJCF!	Sep 3, 2020	Edit / Post a note / Update link / Delete
Lecture 1: Logistics and Intro	Sep 3, 2020	Edit / Post a note / Update file / Delete

General Resources

Manually sort using ≡ Sort on: - -

General Resources	Actions
Demo: Creating a pdf of an ipynb and submitting on Gradescope	Edit, Post a note, Update file, Delete
Review: Probability Theory	Edit, Post a note, Update link, Delete
Review: Partial Derivatives (Section 1 and its exercises)	Edit, Post a note, Update link, Delete
Review: Linear Algebra	Edit, Post a note, Update file, Delete
Syllabus	Edit, Post a note, Update file, Delete

Add Links Add Files

Lab Sessions

Manually sort using ≡ Sort on: - -

Lab Sessions	Date	Actions
Lab 3: Video, Password: dP^tFP2&	Oct 2, 2020	Edit, Post a note, Update link, Delete
Lab 3: Tensorflow introduction code	Oct 2, 2020	Edit, Post a note, Update file, Delete
Lab 3: Tensorflow introduction	Oct 2, 2020	Edit, Post a note, Update file, Delete
Lab 2: Video, Password: 9jNvgU+q	Sep 25, 2020	Edit, Post a note, Update link, Delete
Lab 2: Numpy tutorial	Sep 25, 2020	Edit, Post a note, Update file, Delete
Lab 2: Numpy tutorial code	Sep 25, 2020	Edit, Post a note, Update file, Delete
Lab 1: Notes	Sep 18, 2020	Edit, Post a note, Update file, Delete
Lab 1: Video, Password: $!d9+3^x	Sep 18, 2020	Edit, Post a note, Update link, Delete
Lab 1: Preliminaries	Sep 18, 2020	Edit, Post a note, Update file, Delete

Add Links Add Files

Video Demo

- Installing Tensorflow
- Converting ipynb to pdf
- Submitting pdf on gradescope

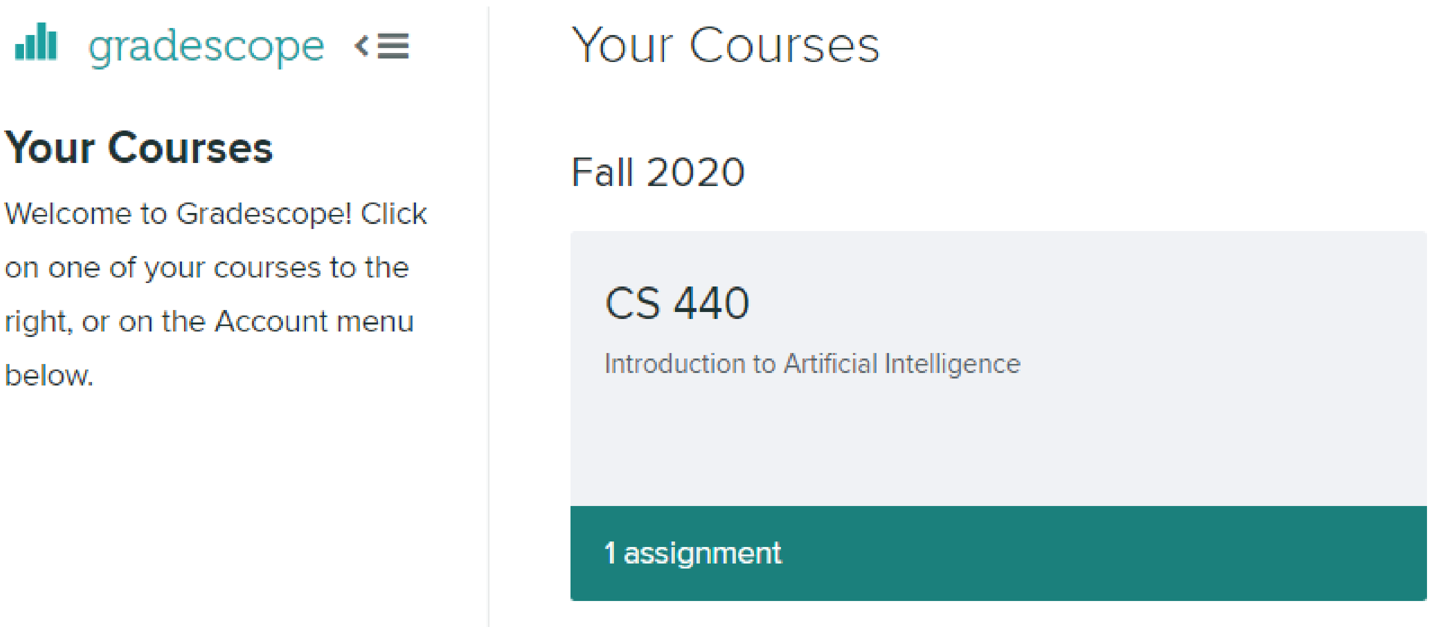
https://www.youtube.com/watch?v=gxR012ZAR4o

Gradescope

- Please join the Gradescope with your BU email

https://www.gradescope.com/courses/186182%E2%80%A9Entry

- Entry Code: 95Y7JG



Dry Run

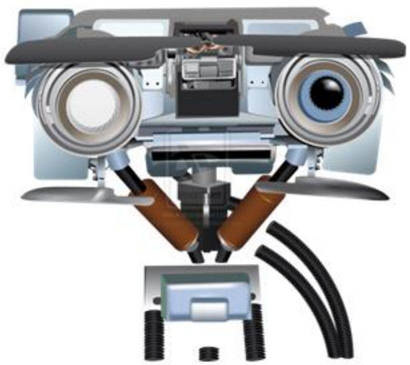
- Create a pdf from the current notebook with no solutions
- Submit it now in gradescope as a dry run
- Do not wait until you have completed solving the assignment to try out the submission system
- This will give you ample time to ask any questions you have about the submission process before the deadline and will make your future submissions smooth

Important!

- ***Keep a copy of your ipynb with unchanged last date of modification***
  - Graders might need it, and would be requesting it by email cc ing me

- **Friday's Lab:**
  brief demo + office hour for help with PS1

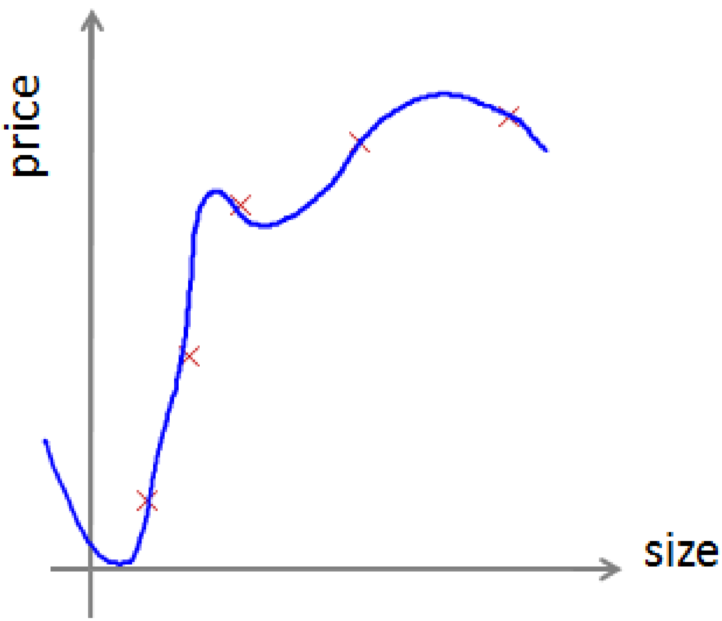
- We encourage you to make best use of Piazza and Office Hours



Model Selection

Training/Validation/Test Sets

Overfitting example



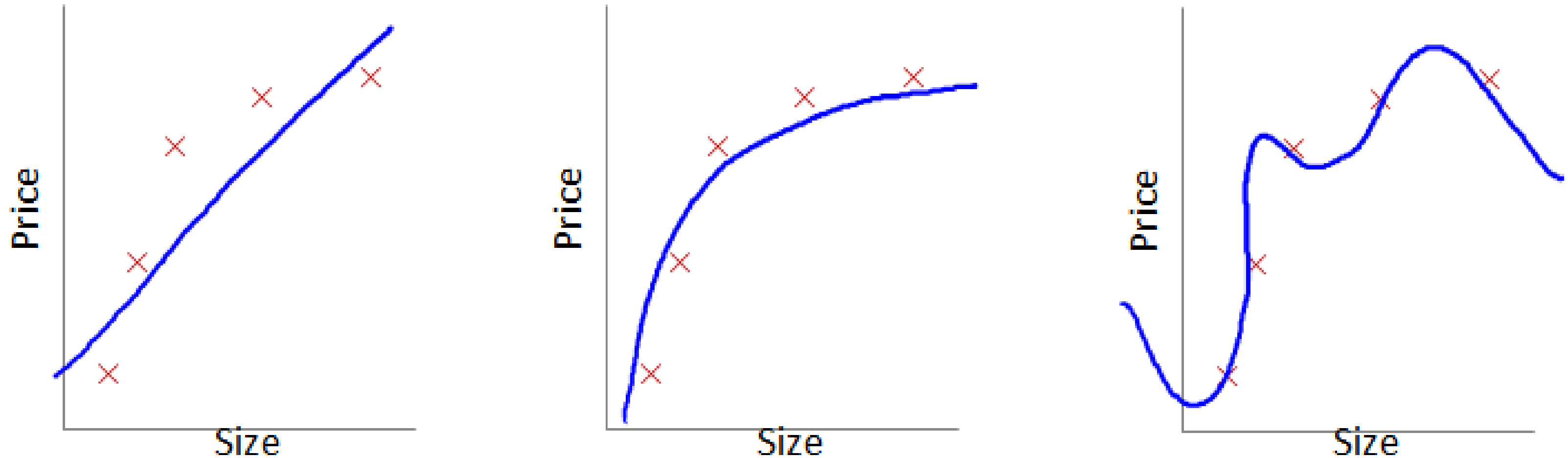
$$h_\theta(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$$

Once parameters $(\theta_0, \theta_1, \ldots, \theta_4)$ were fit to some set of data (training set), the error of the parameters as measured on that data (the training error $J(\theta)$) is likely to be lower than the actual generalization error.

One solution is to regularize, but how can we choose the regularization weight λ?

Choosing weight λ

$$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$$



$\lambda = 100$

High bias (underfit)

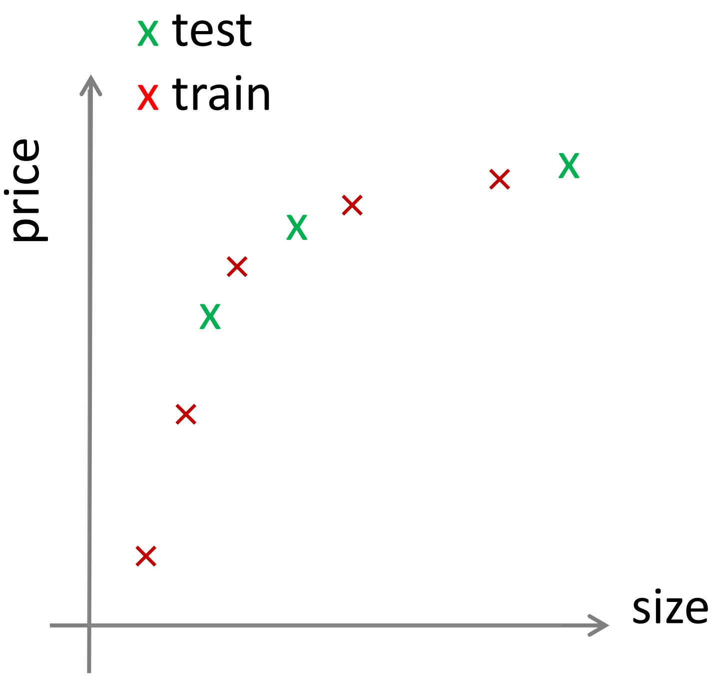
$\lambda = 1$

"Just right"

$\lambda = 0.01$

High variance (overfit)

Model selection



$$h_\theta(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$$

Hyperparameters (e.g., degree of polynomial, regularization weight, learning rate) must be selected prior to training.

How to choose them?

Try several values, choose one with the lowest test error?

Problem: test error is likely an overly optimistic estimate of generalization error because we “cheat” by fitting the hyperparameter to the actual test examples.

Train/Validation/Test Sets

	Size	Price
train	2104	400
	1600	330
	2400	369
	1416	232
	3000	540
validation	1985	300
	1534	315
	1427	199
test	1380	212
	1494	243

Solution: split data into three sets.

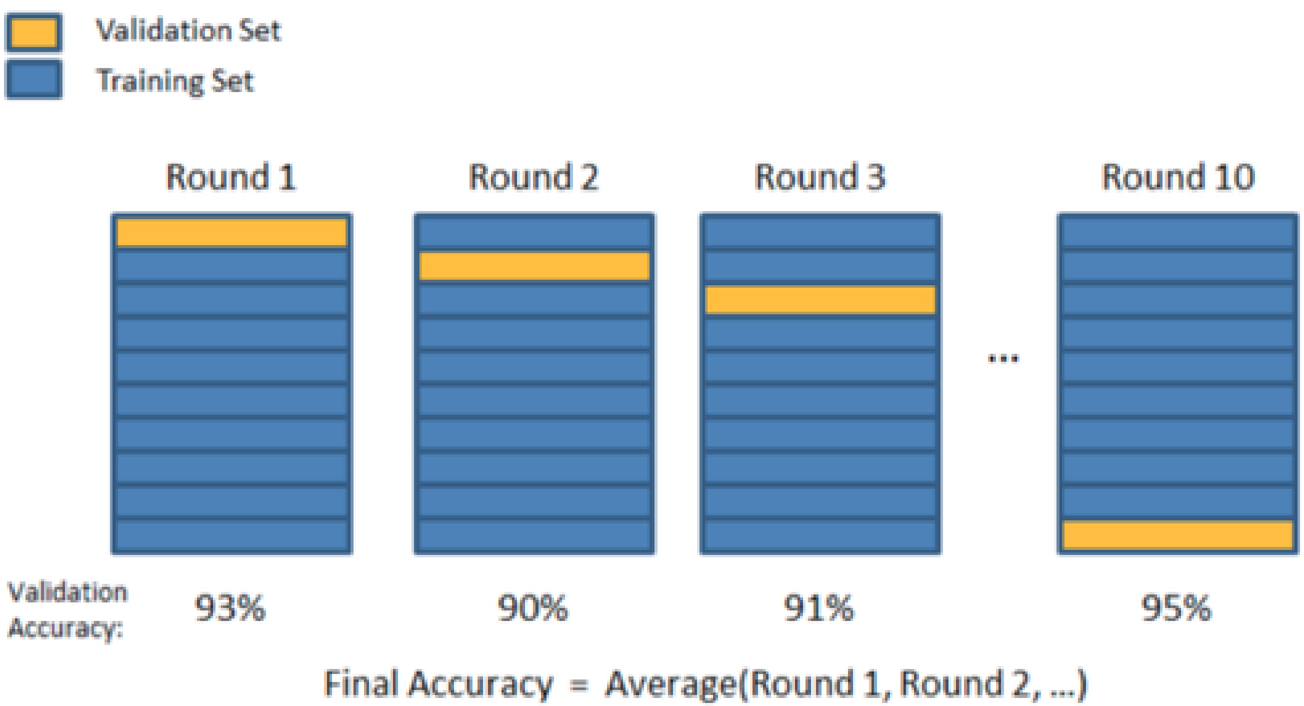
For each value of a hyperparameter, train on the train set, evaluate learned parameters on the validation set.

Pick the model with the hyper parameter that achieved the lowest validation error.

Report this model's test set error.

N-Fold Cross Validation

- What if we don't have enough data for train/test/validation sets?
- Solution: use N-fold cross validation.
- Split training set into train/validation sets N times.
- Report average predictions over N val sets, e.g. N=10:



Confusion Matrix

- A performance measurement for machine learning classification problem where output can be two or more classes.

- **True Positive:**
  predicted positive and it's true

  True Negative:
  predicted negative and it's true

  False Positive:
  predicted positive and it's false

  False Negative:
  predicted negative and it's false

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

Recall

- Out of all the positive classes, how much we predicted correctly. It should be high as possible.

$$\textbf{\textit{Recall}} = \frac{TP}{TP + FN}$$

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

Precision

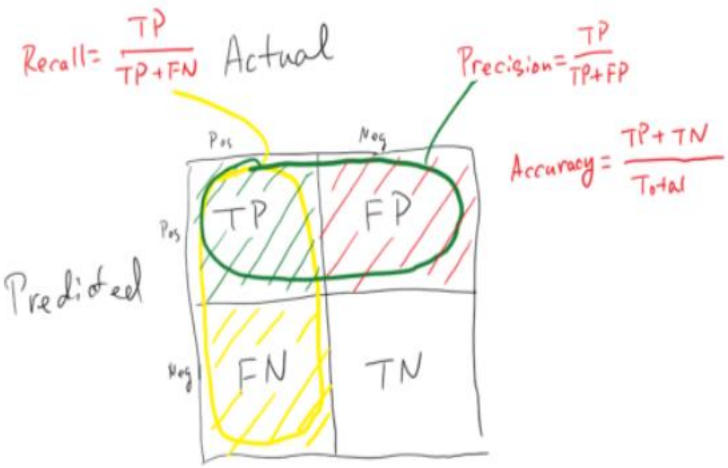
- Out of all the positive classes we have predicted, how many are actually positive.

$$\textit{Precision} = \frac{TP}{TP + FP}$$

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Example

y	y pred	output for threshold 0.6	Recall	Precision	Accuracy
0	0.5	0	1/2	2/3	4/7
1	0.9	1			
0	0.7	1			
1	0.7	1			
1	0.3	0			
0	0.4	0			
1	0.5	0			



F-measure

- F-score helps to measure Recall and Precision at the same time.

$$F\text{-}measure = \frac{2 * Recall * Precision}{Recall + Precision}$$

https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

Neural Networks VI

Explainability

Importance of *Explainability*

- An important action to be detected in the vision systems of autonomous vehicles is: *Pedestrian Crossing*

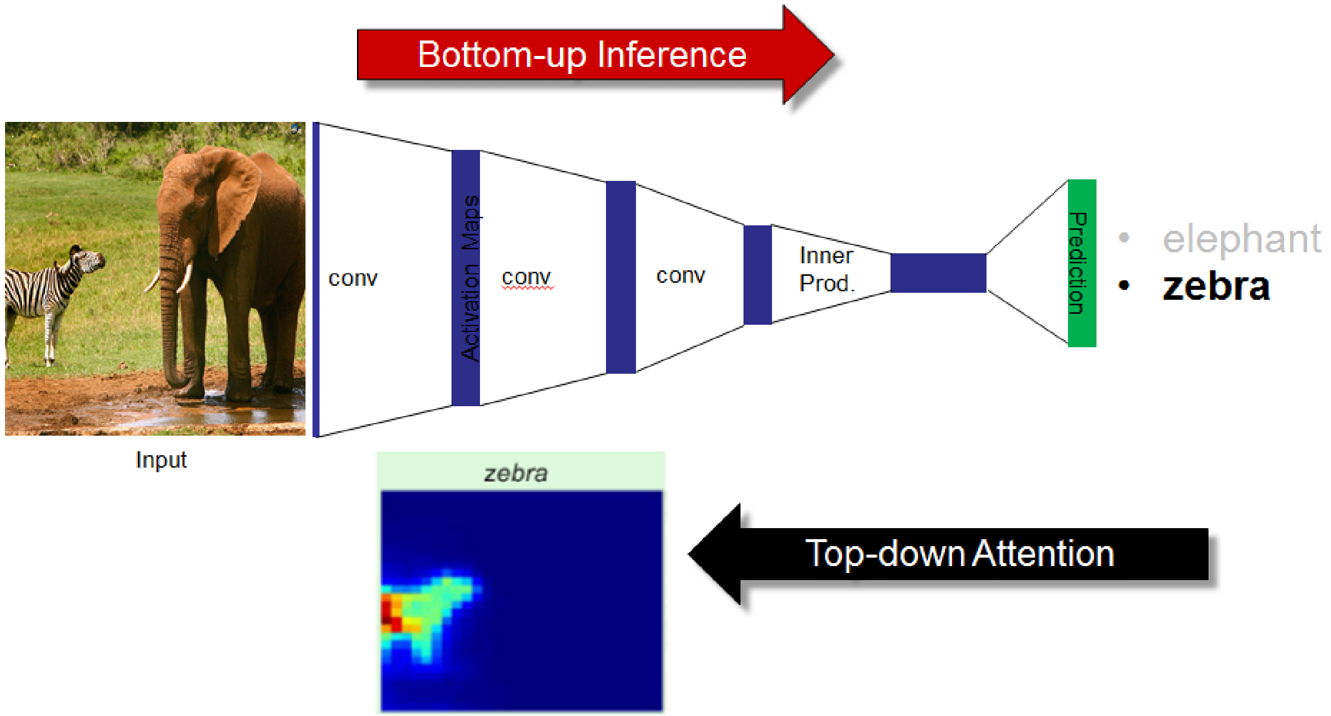
Importance of *Explainability*

Sample Misclassification

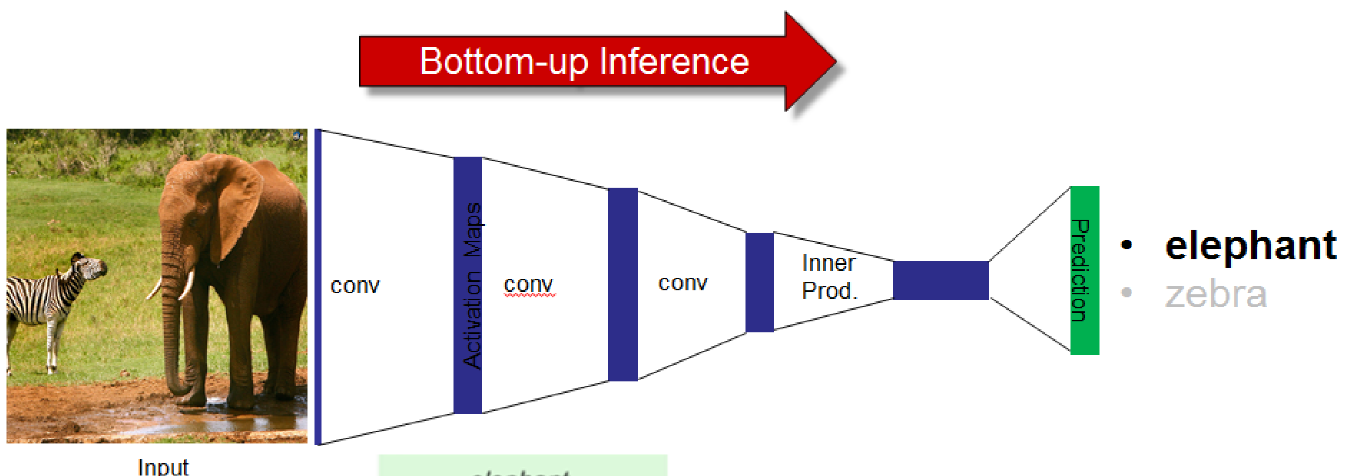
Ground Truth:
BabyCrawling

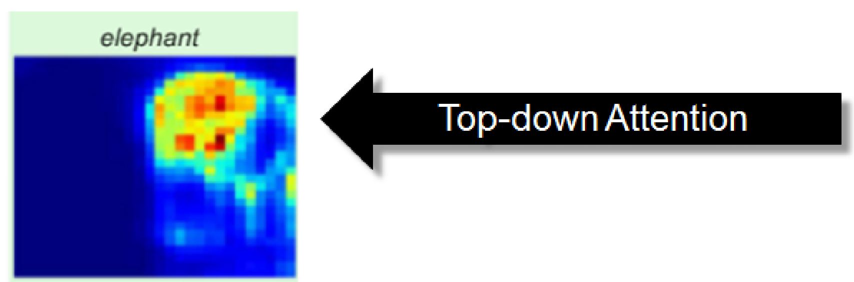
Classified as:
Pushups

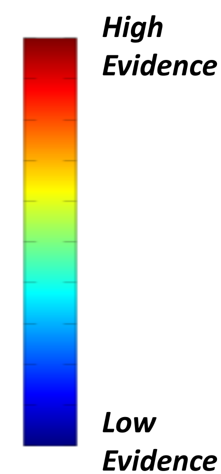
Spatial Grounding



Spatial Grounding

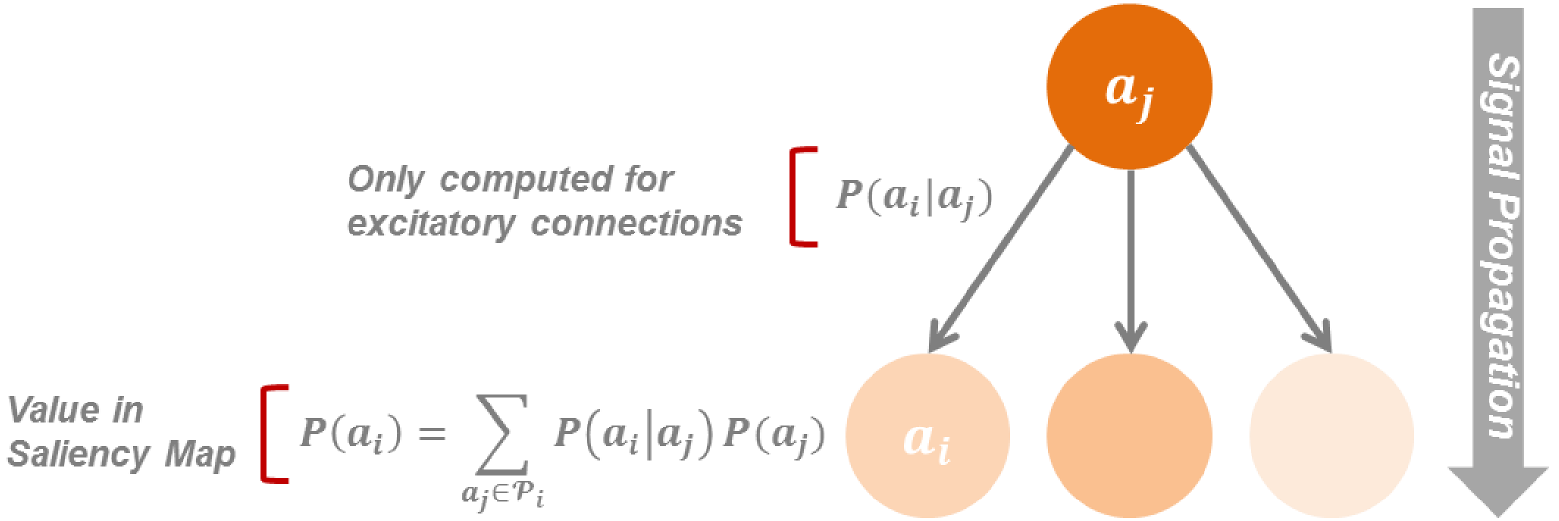






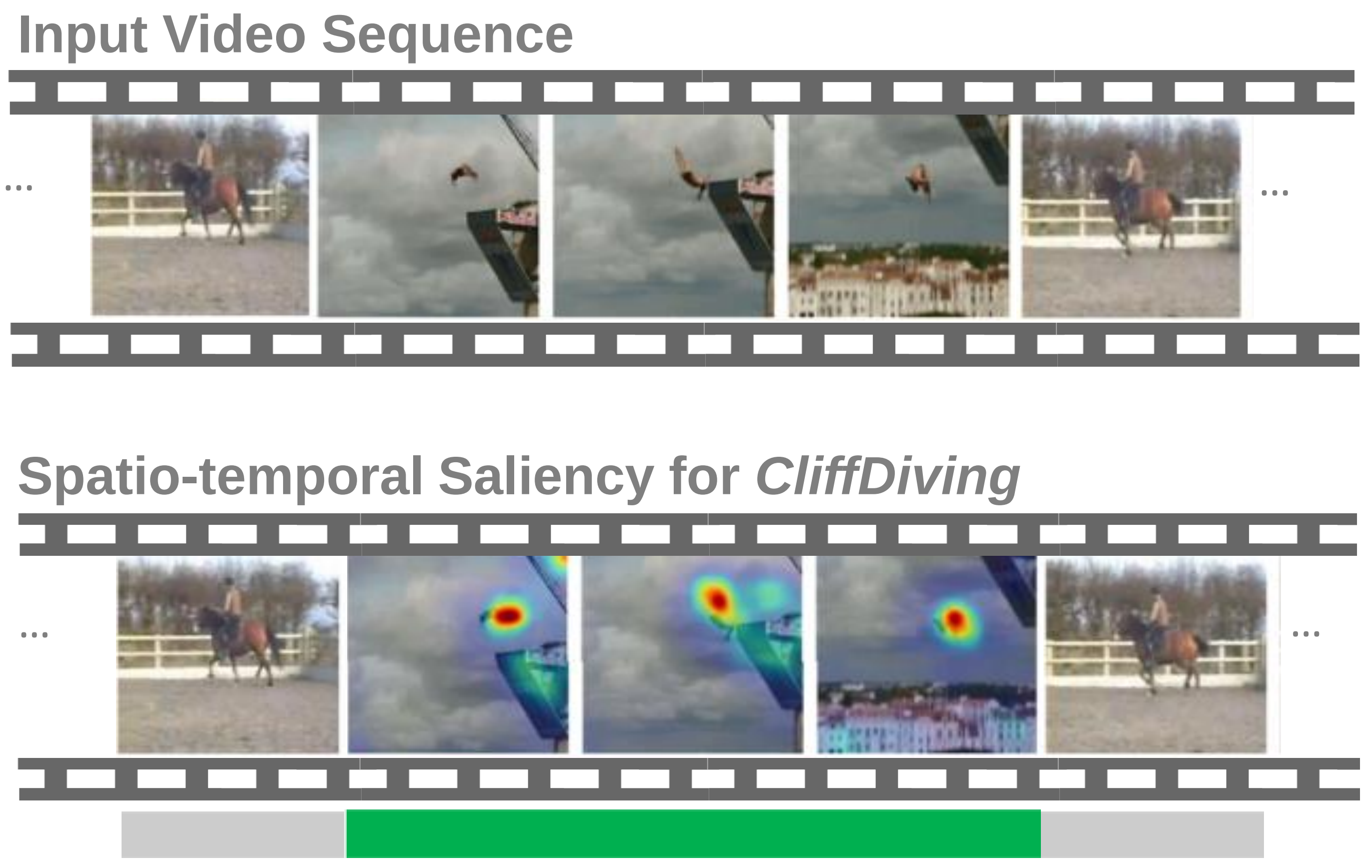
- *Evidence*
- *Grounding*
- *Saliency*

Excitation Backprop (EB)



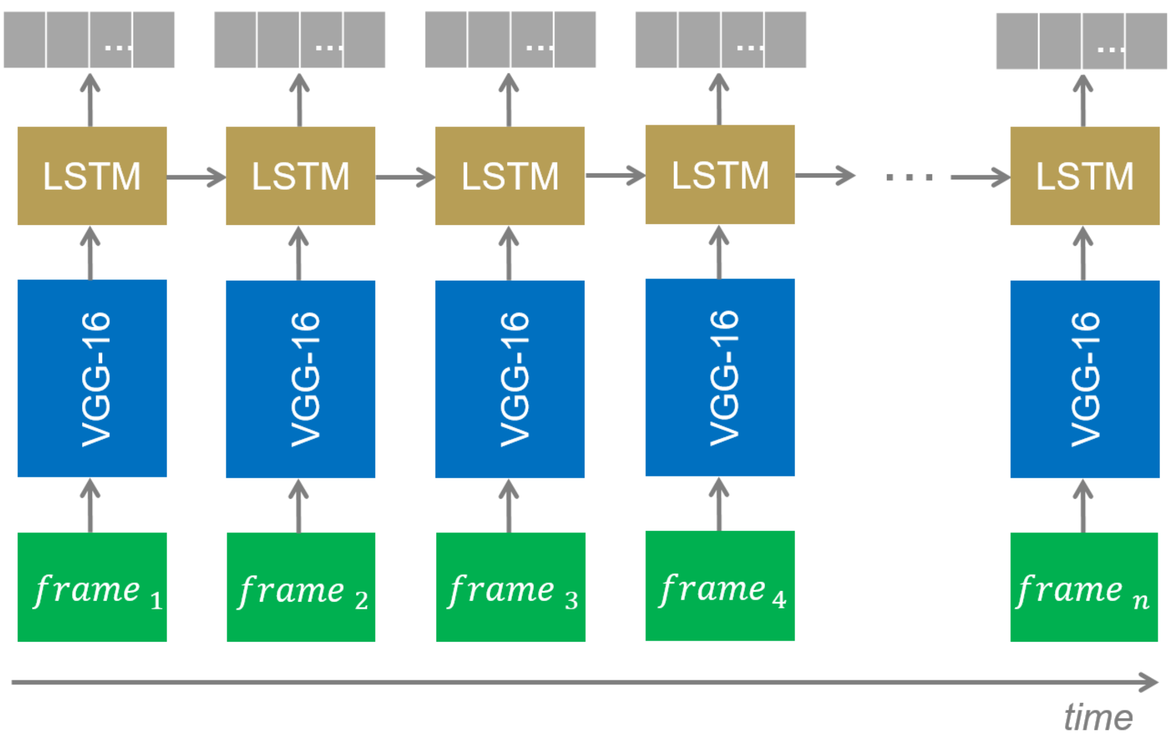
[Jianming Zhang, Zhe Lin, Jonathan Brandt, Xiaohui Shen, Stan Sclaroff. "*Top-down Neural Attention by Excitation Backprop., ECCV'16*]

Spatiotemporal Grounding

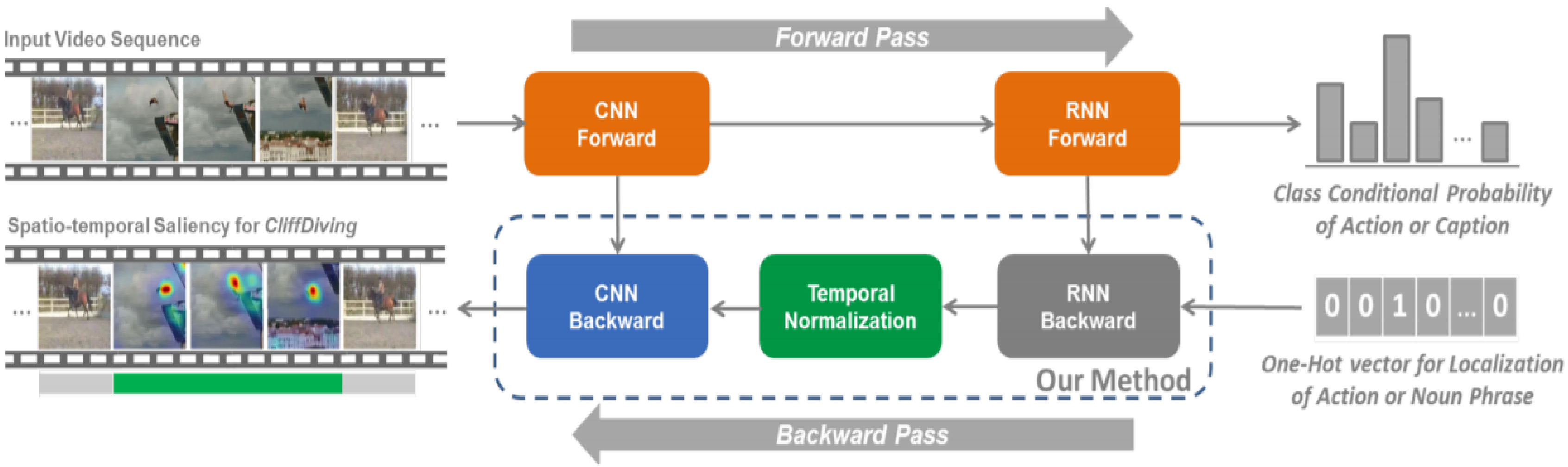


Architecture: Forward Pass

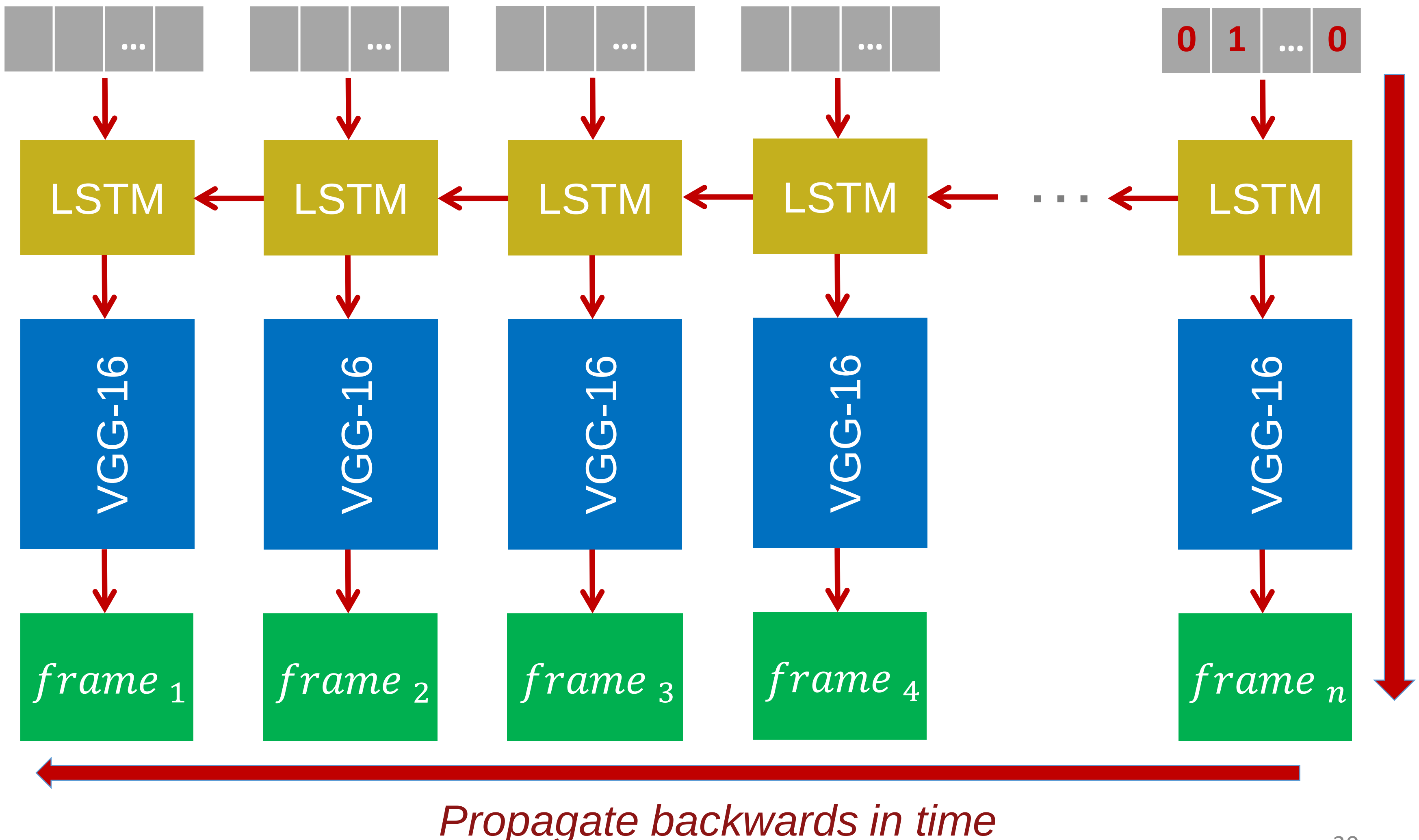
- CNN-LSTM is trained for the action recognition task.
- Resulting grounding is weakly-supervised.



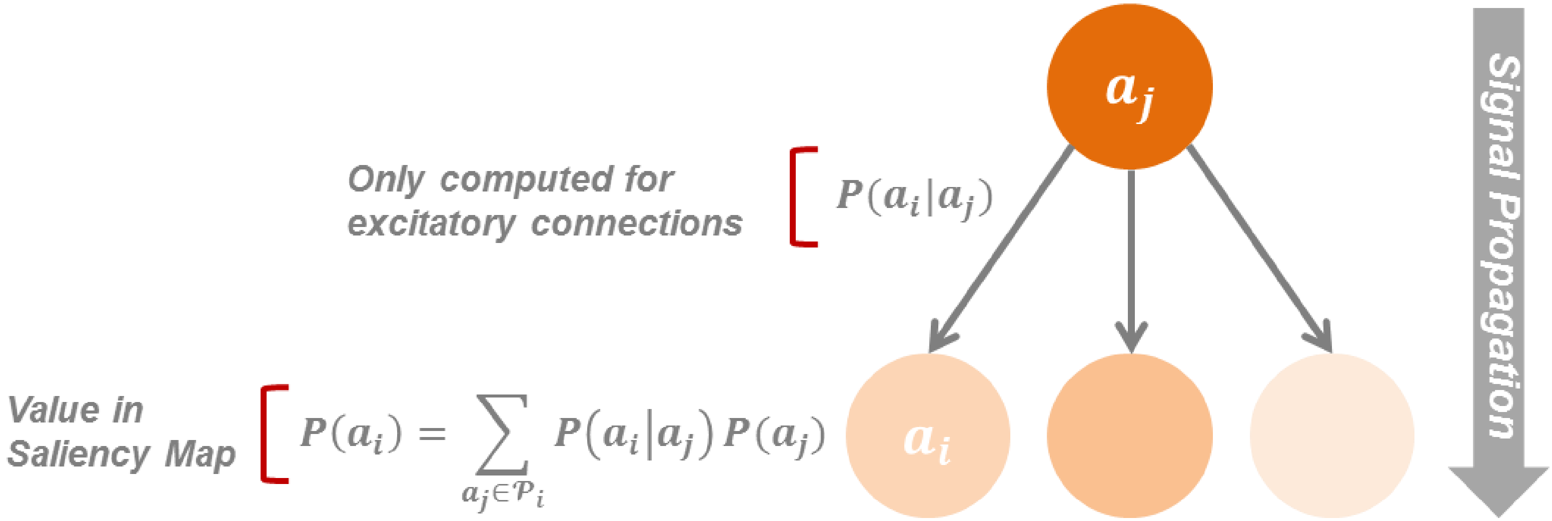
Excitation Backprop in RNNs



Architecture: Backward Grounding Pass



Excitation Backprop (EB)



[Jianming Zhang, Zhe Lin, Jonathan Brandt, Xiaohui Shen, Stan Sclaroff. "*Top-down Neural Attention by Excitation Backprop., ECCV'16*]

Applications

- Action Detection (*videos*)

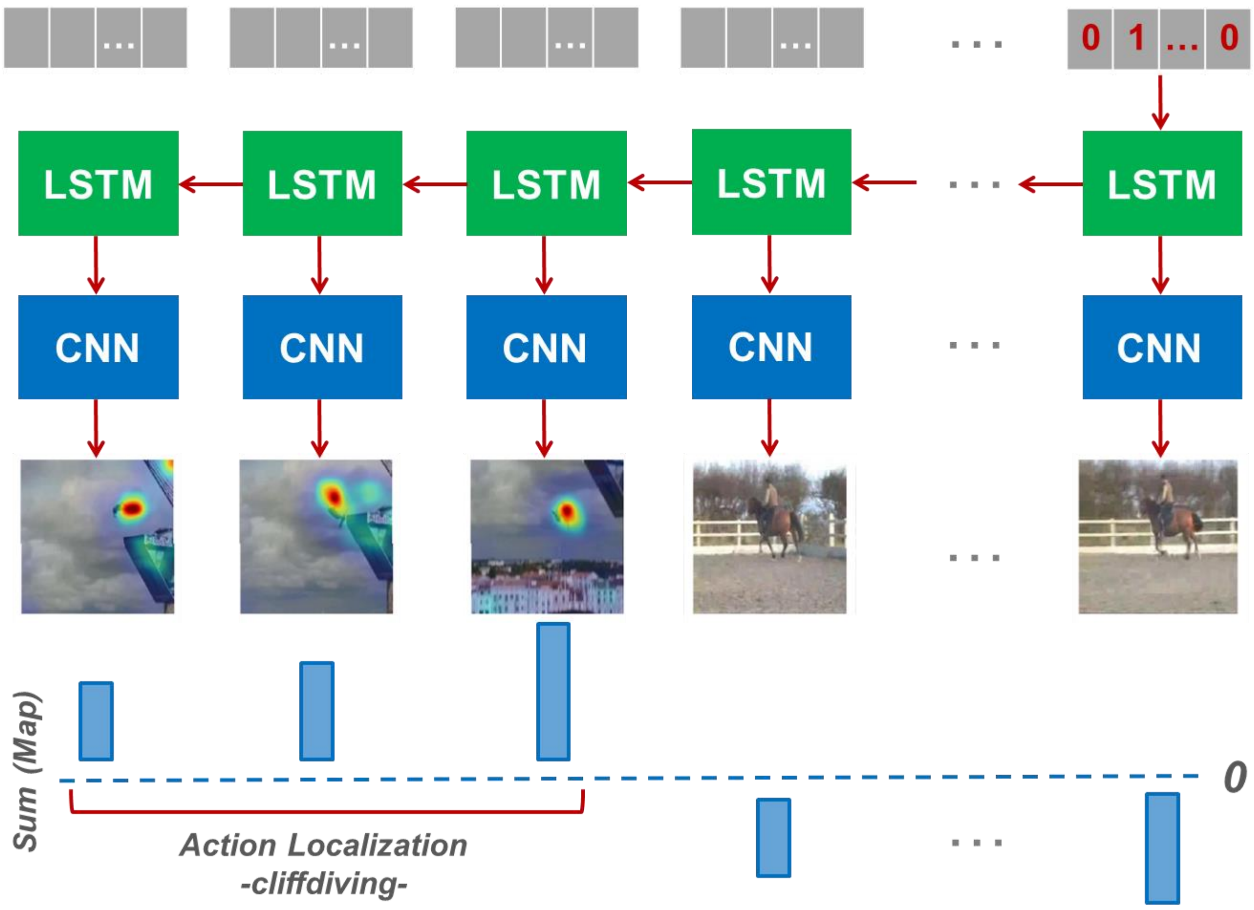
- Caption Grounding (*images, videos*)

- Reflecting the Abstraction Capability of Models

Applications

- Action Detection (*videos*)
- Caption Grounding (*images, videos*)
- Reflecting the Abstraction Capability of Models

Spatiotemporal Action Detection



Applications

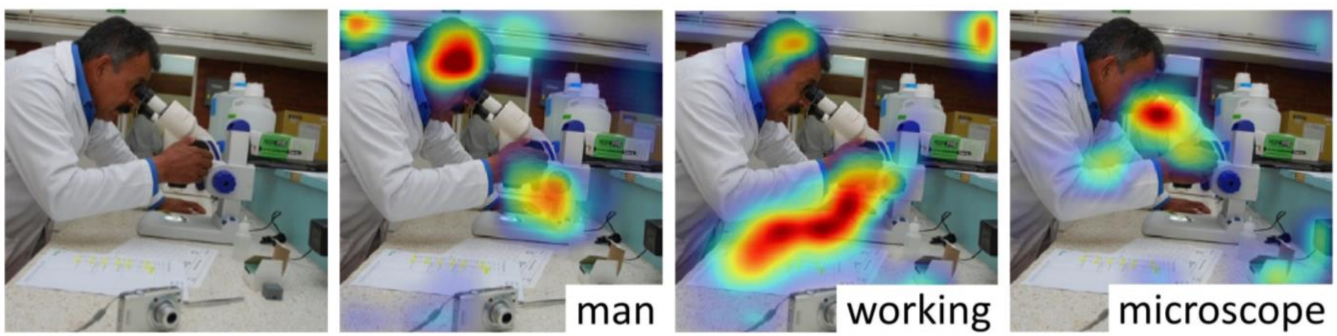
- Action Detection (*videos*)
- Caption Grounding (*images, videos*)
- Reflecting the Abstraction Capability of Models

Flicker30kEntities Dataset: Grounding Words of an Image Caption

image caption: *A man in a lab coat is working on a microscope.*

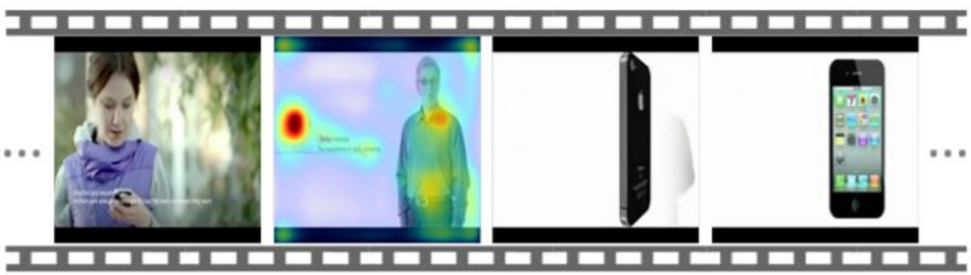
Flicker30kEntities Dataset: Grounding Words of an Image Caption

image caption: *A man in a lab coat is working on a microscope.*



MSRVTT Dataset: Grounding Words of a Video Caption

video caption: *"A man is talking about a phone"*



(a) grounding of the word *man*

(b) grounding of the word *phone*

Applications

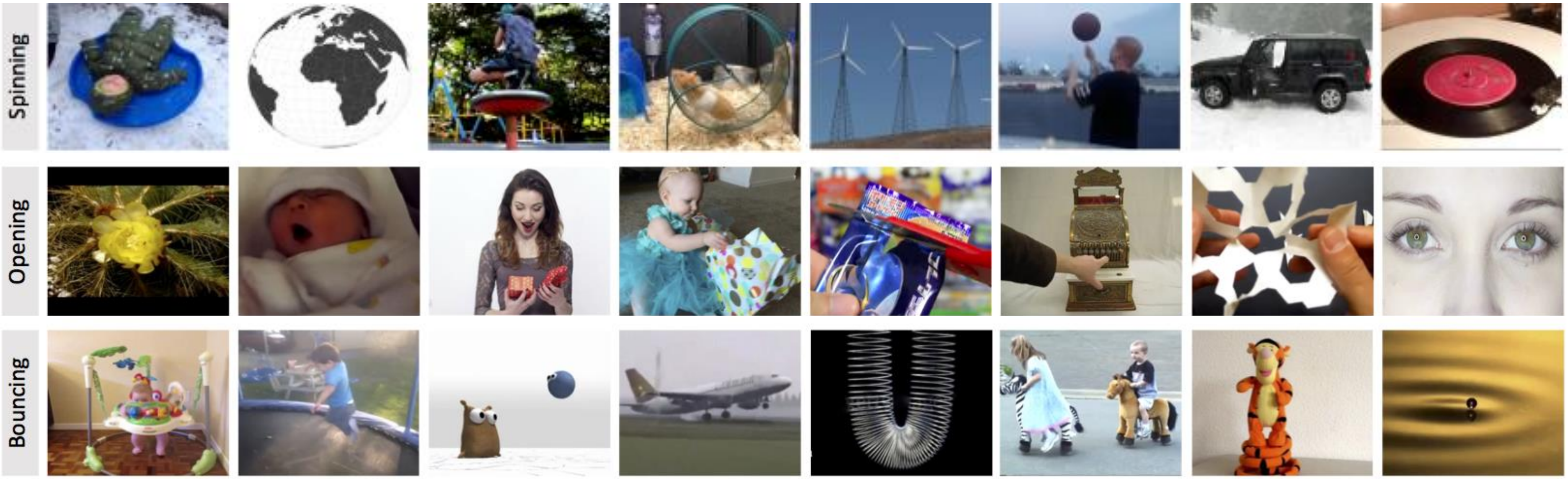
- Action Detection (*videos*)
- Caption Grounding (*images, videos*)
- Reflecting the Abstraction Capability of Models

Reflecting the Abstraction Capability of Models

- Moments in Time Dataset

  M. Monfort, A. Andonian, B. Zhou, K. Ramakrishnan, S. A. Bargal, T. Yan, L. Brown, Q. Fan, D. Gutfruend, C. Vondrick, A. Oliva. “Moments in Time Dataset: one million videos for event understanding.” *TPAMI*, 2019.

- Videos of abstract dynamical events performed by various actors.



Moments in Time Dataset

- Typically, classification accuracy is reported to summarize the recognition capability of models.
- However, classification accuracy alone is not representative as to whether the models are really modeling this diversity of actors.
- A classifier may be incorrectly classifying a whole subset of cases/actors.

Moments in Time Dataset

- Class: *Opening*