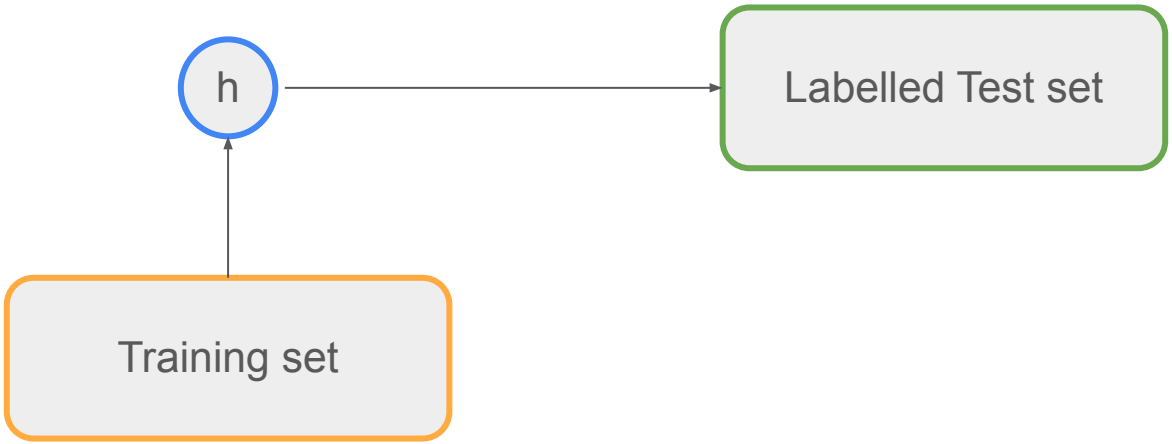
Predicting Classifier Accuracy under Prior Probability Shift

Lorenzo Volpi

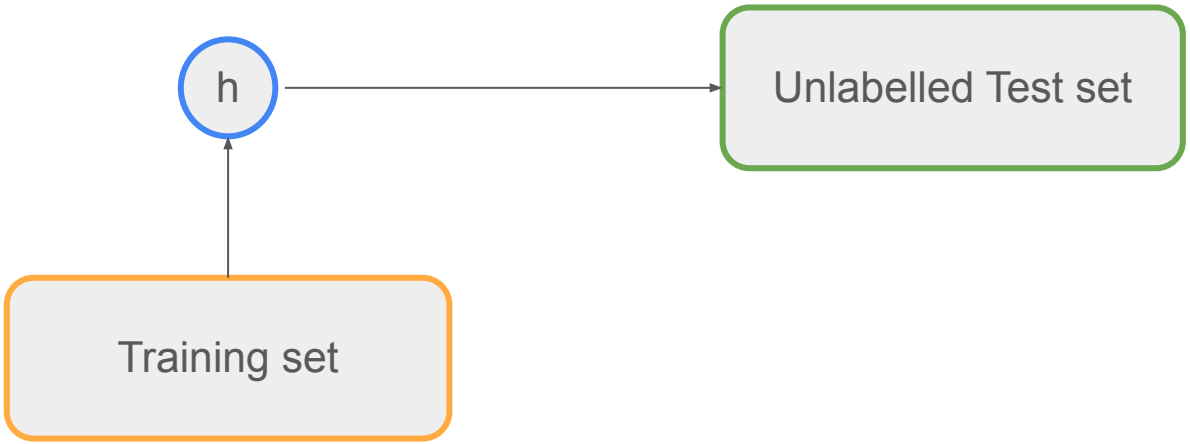
24/05/2024

Classifier Accuracy Evaluation



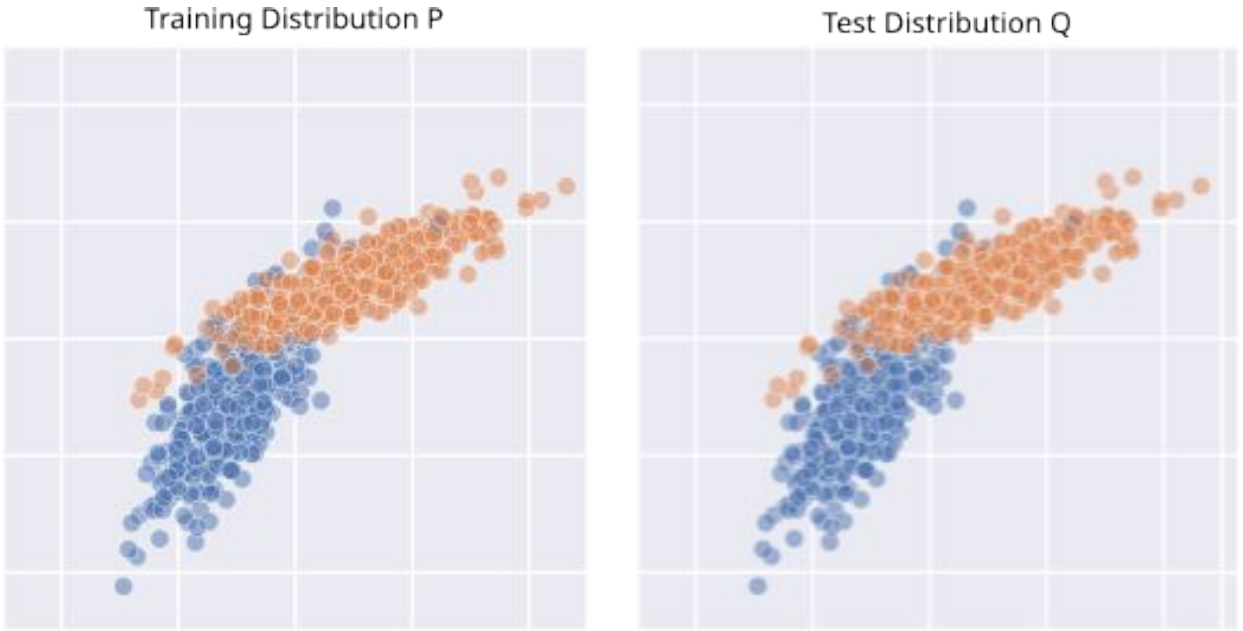
- “Easy” to achieve
- Many datasets and benchmarks

Classifier Accuracy Prediction (**CAP**)



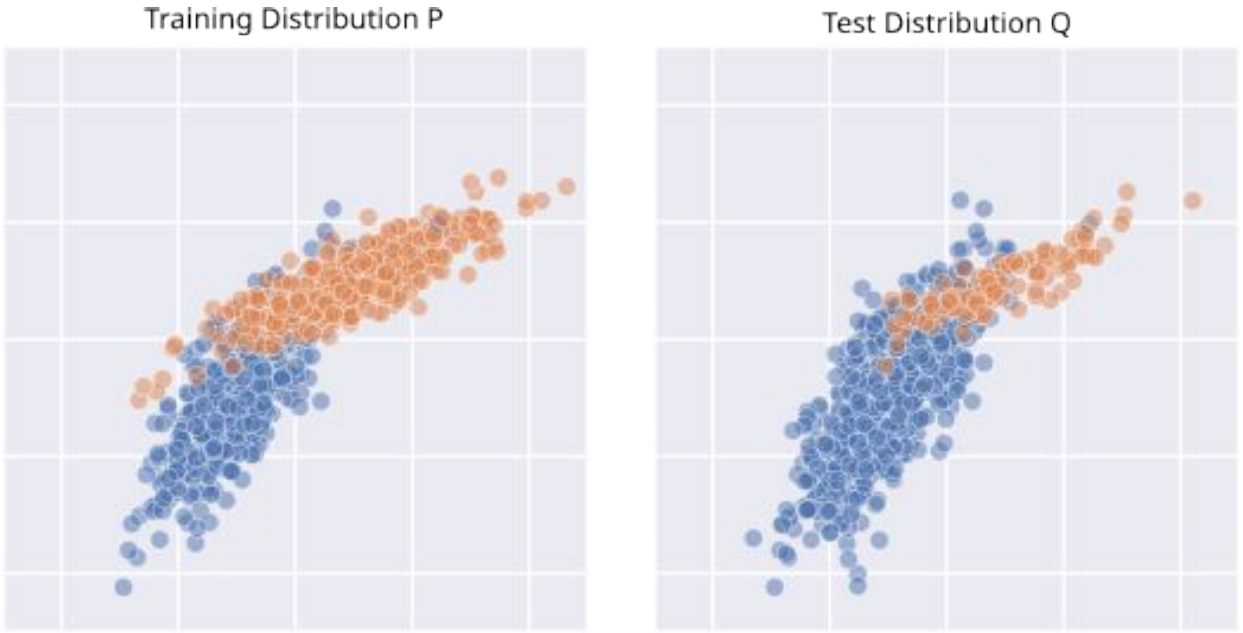
- Cannot be directly estimated —Prediction
- K-fold Cross Validation
  - Not good under Dataset Shift

Independent and Identically Distributed (IID)



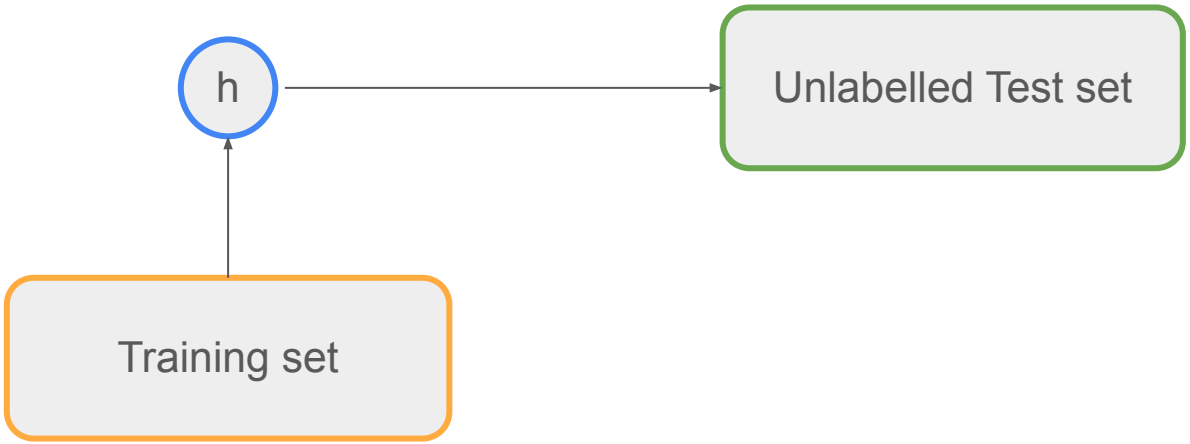
$$P(X,y) = Q(X,y)$$

Prior Probability Shift (PPS)



$$P(y) \neq Q(y)$$

Classifier Accuracy Prediction (**CAP**)



- Cannot be directly estimated — ~~Prediction~~
- K-fold Cross Validation
  - Not good under Dataset Shift
- **Quantification!**

Classification

If you like original gut wrenching laughter you will like this movie. If you are young or old then you will love this movie, hell even my mom liked it.

Great Camp!!!

positive

The plot is about the death of little children. Hopper is the one who has to investigate the killings. During the movie it appears that he has some troubles with his daughter. In the end the serial killer get caught. That's it. But before you find out who dunnit, you have to see some terrible acting by all of the actors. It is unbelievable how bad these actors are, including Hopper. I could go on like this but that to much of a waste of my time. Just don't watch the movie. I've warned you.

negative

Quantification

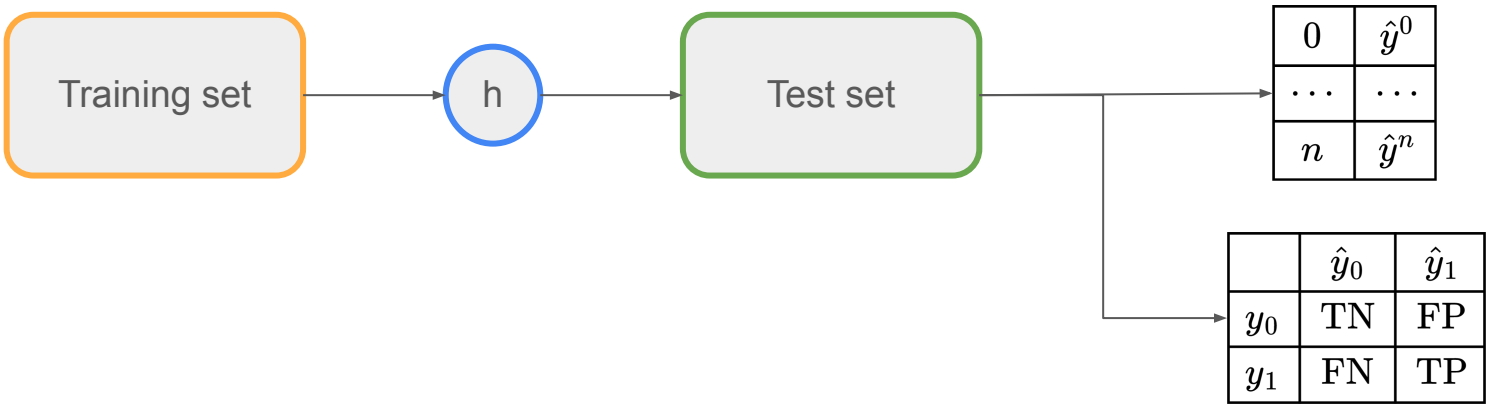
- How many positive reviews in the whole dataset?
- Predict prevalence values for each class of the dataset

Quantification

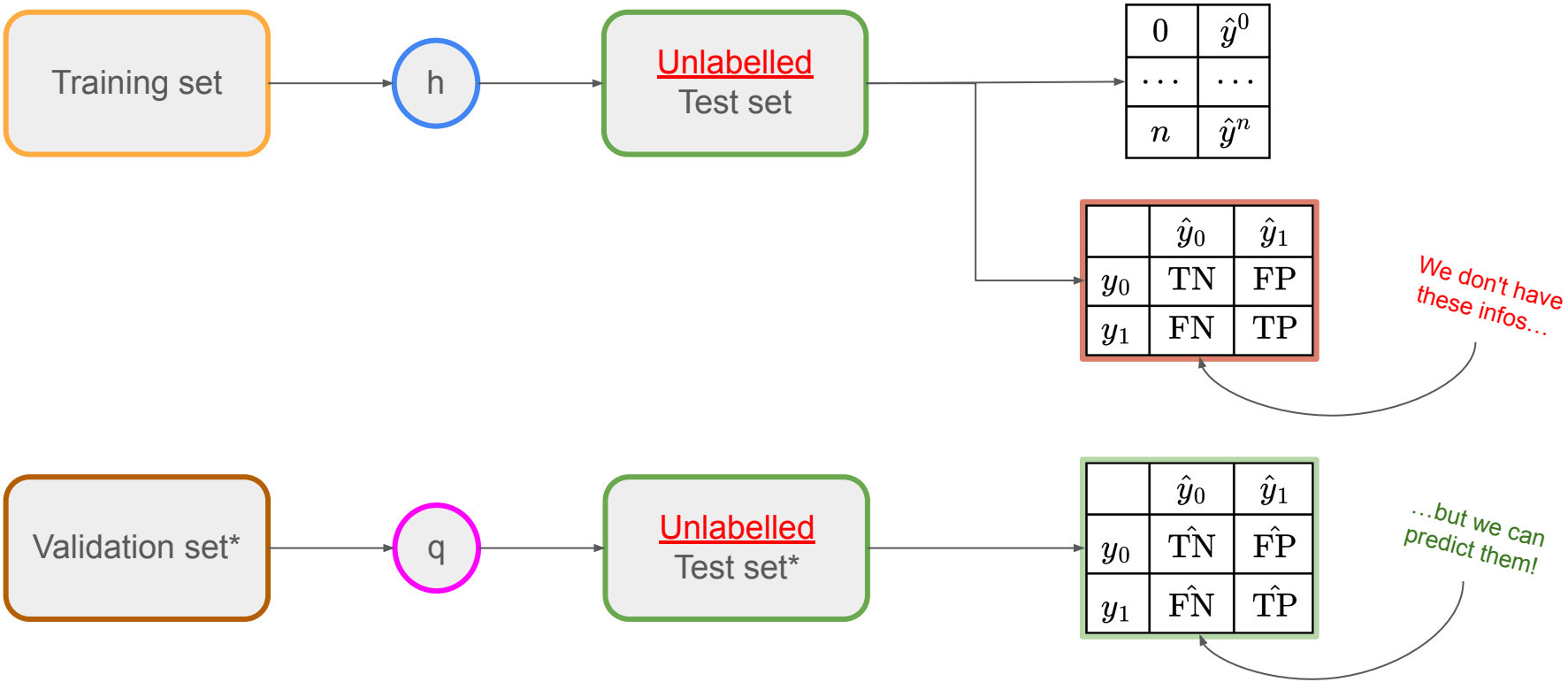
- How many positive reviews in the whole dataset?
- Predict prevalence values for each class of the dataset
- Quantification algorithms:
  - Classify and Count (CC)
  - PACC
  - SLD
  - KDEy

	negative	positive	mae
true prev.	0.9800008	0.019992	/
CC	0.872677	0.127323	0.107330
SLD	0.999804	0.000196	**0.019796**

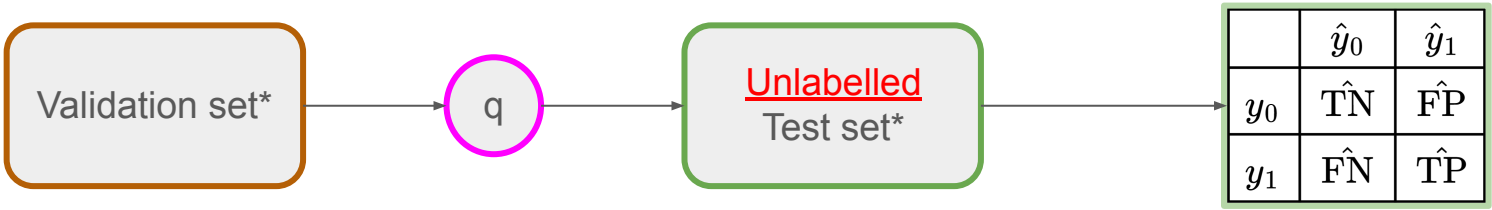
Contingency Table



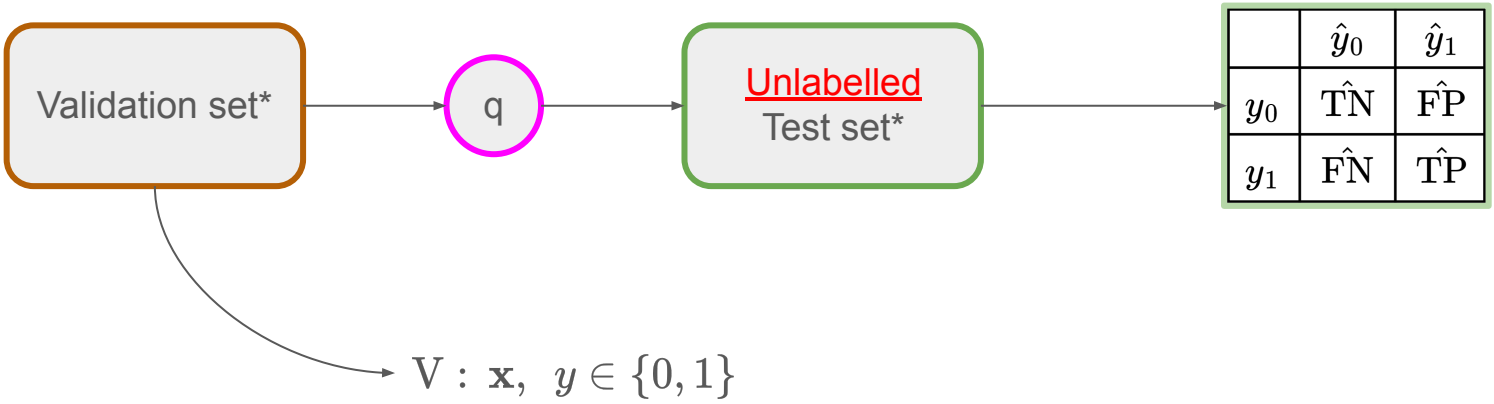
Predicting the Contingency Table



Predicting the Contingency Table

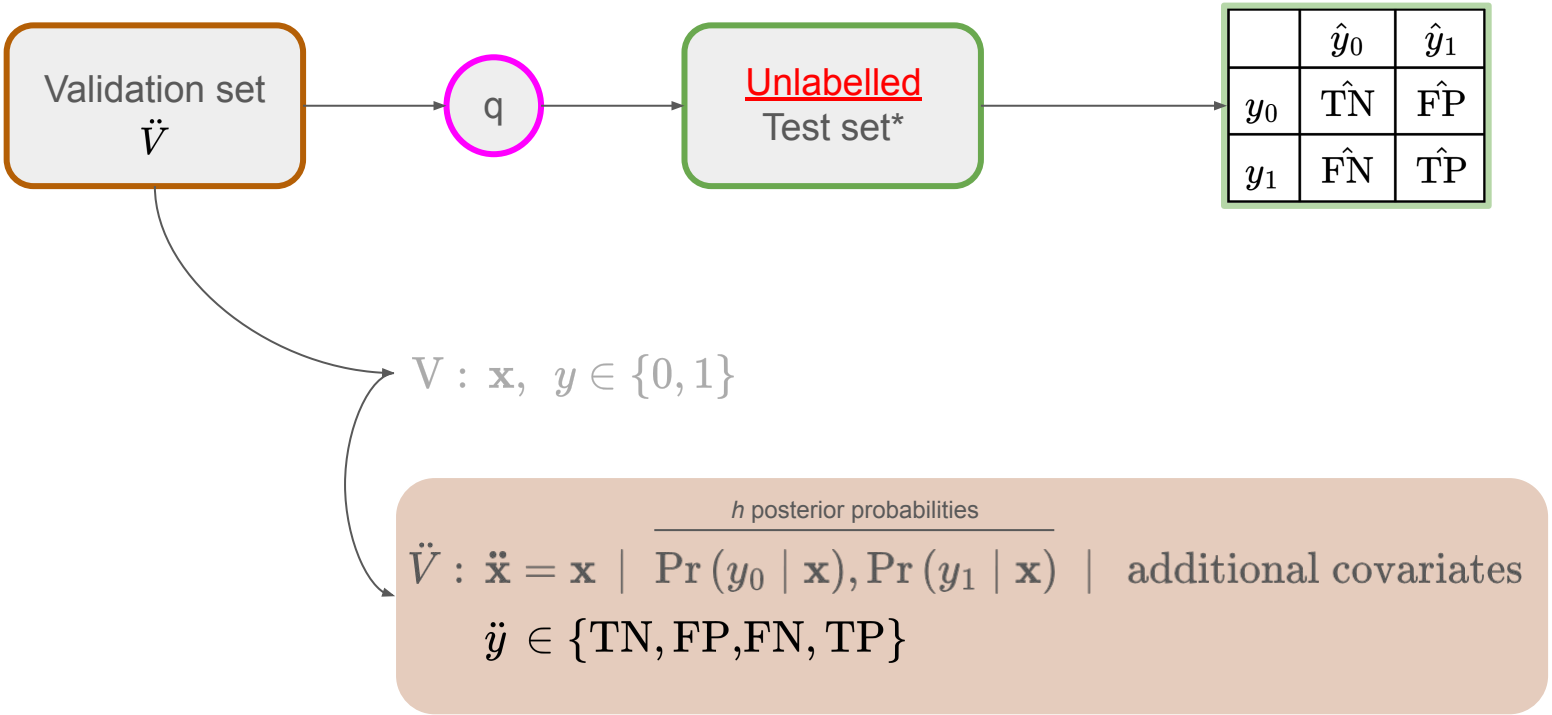


Predicting the Contingency Table

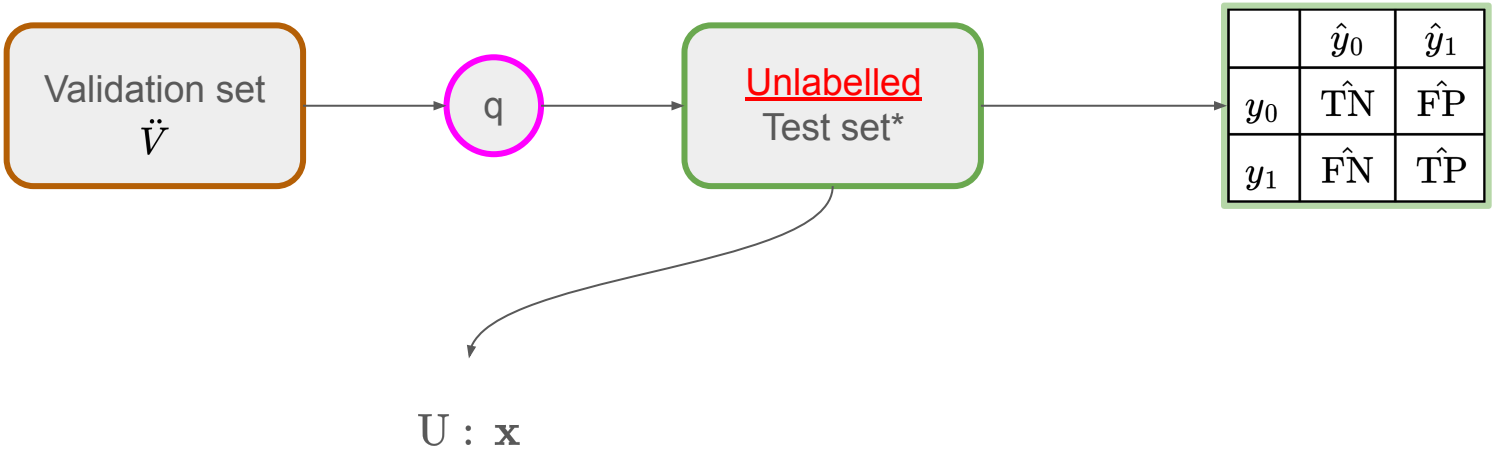


	$\hat{y}_0$	$\hat{y}_1$
y_0	$\hat{\mathrm{TN}}$	$\hat{\mathrm{FP}}$
y_1	$\hat{\mathrm{FN}}$	$\hat{\mathrm{TP}}$

Predicting the Contingency Table

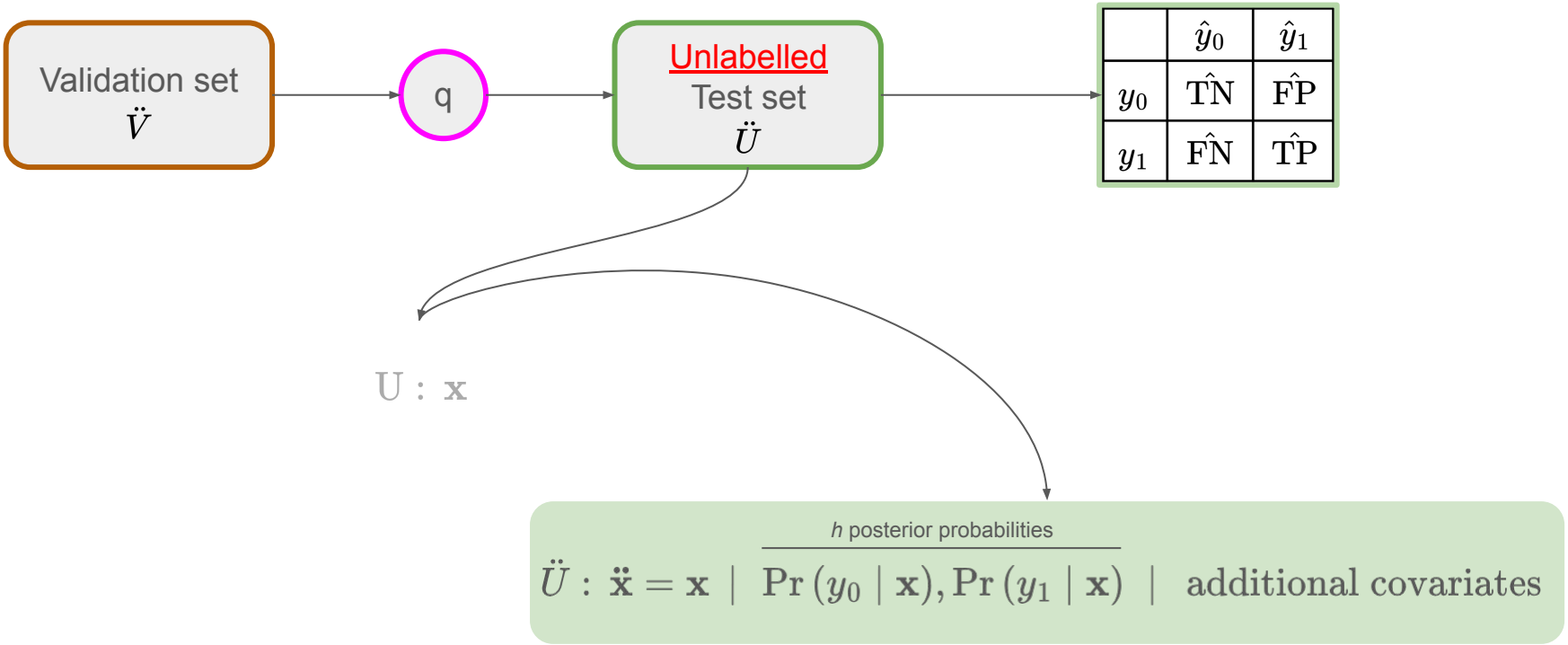


Predicting the Contingency Table

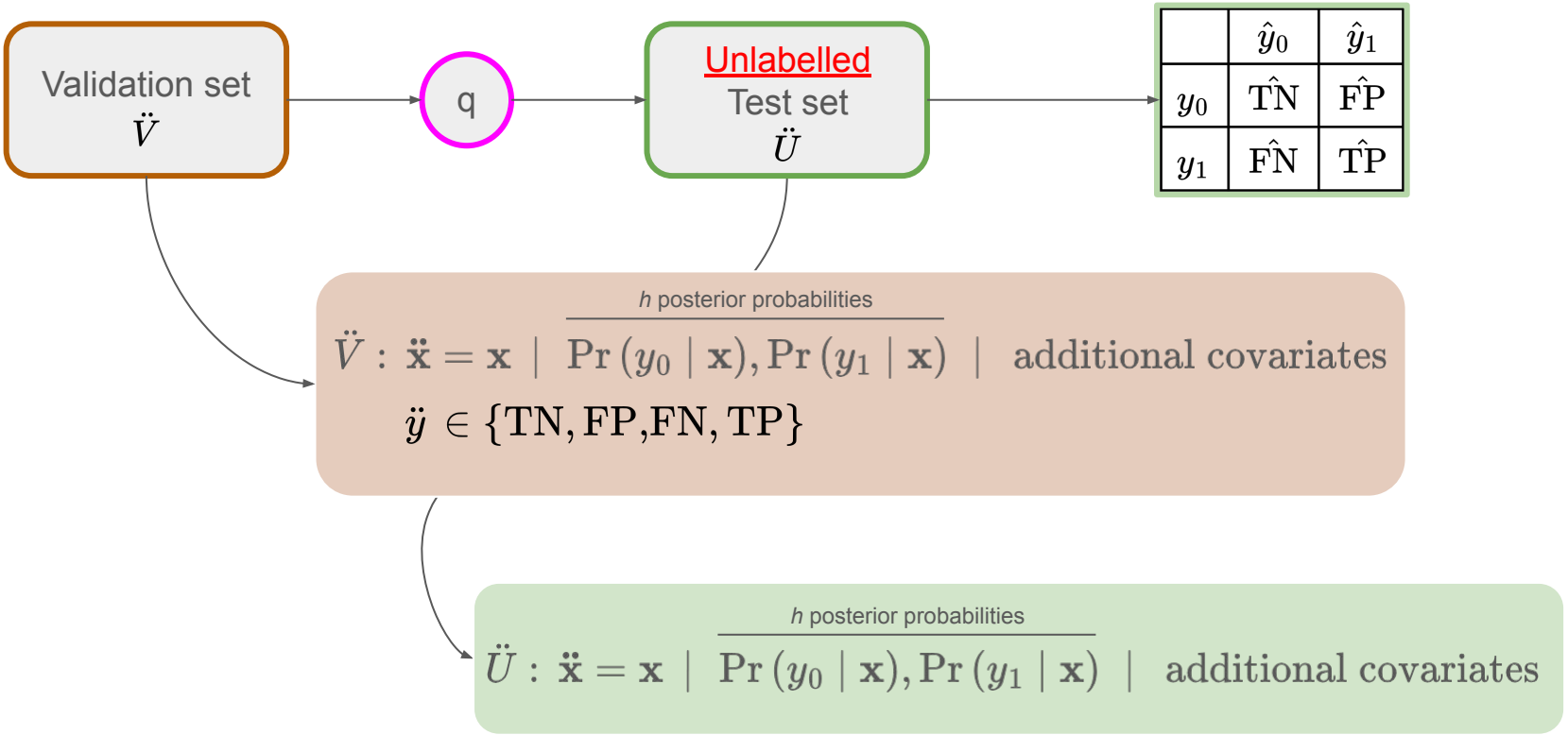


	$\hat{y}_0$	$\hat{y}_1$
y_0	$\hat{\mathrm{TN}}$	$\hat{\mathrm{FP}}$
y_1	$\hat{\mathrm{FN}}$	$\hat{\mathrm{TP}}$

Predicting the Contingency Table



Predicting the Contingency Table



Additional Covariates

max conf

$$\mathrm{MC}(\mathbf{p}) = \max_{\{i \in 1,\ldots,n\}} p_i$$

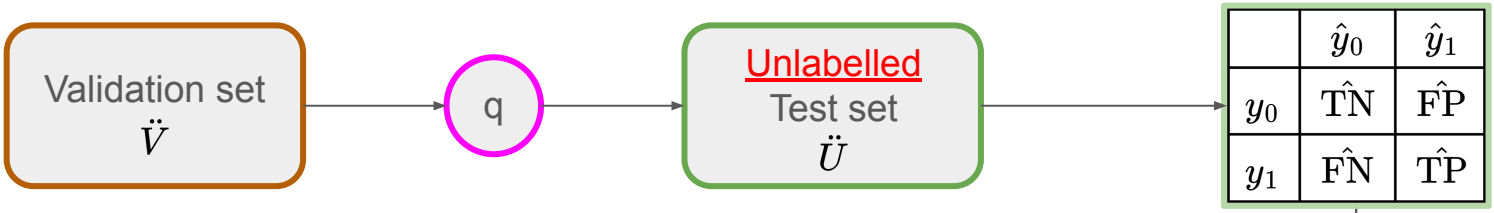
negative entropy

$$\mathrm{NE}(\mathbf{p}) = \sum_{i=1}^{n} p_i \log\ p_i$$

max inverse softmax

$$\mathrm{MIS}(\mathbf{p}) = \max_{i \in \{1,\ldots,n\}} \left(\log\ p_i - \frac{1}{n} \sum_{j=1}^{n} \log\ p_j \right)$$

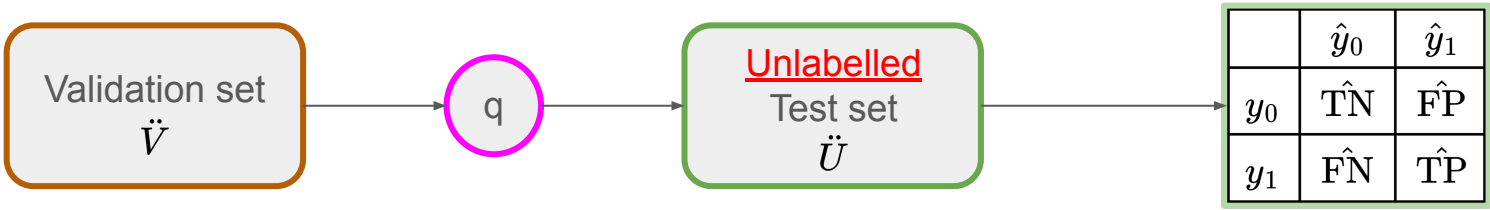
Accuracy Measures



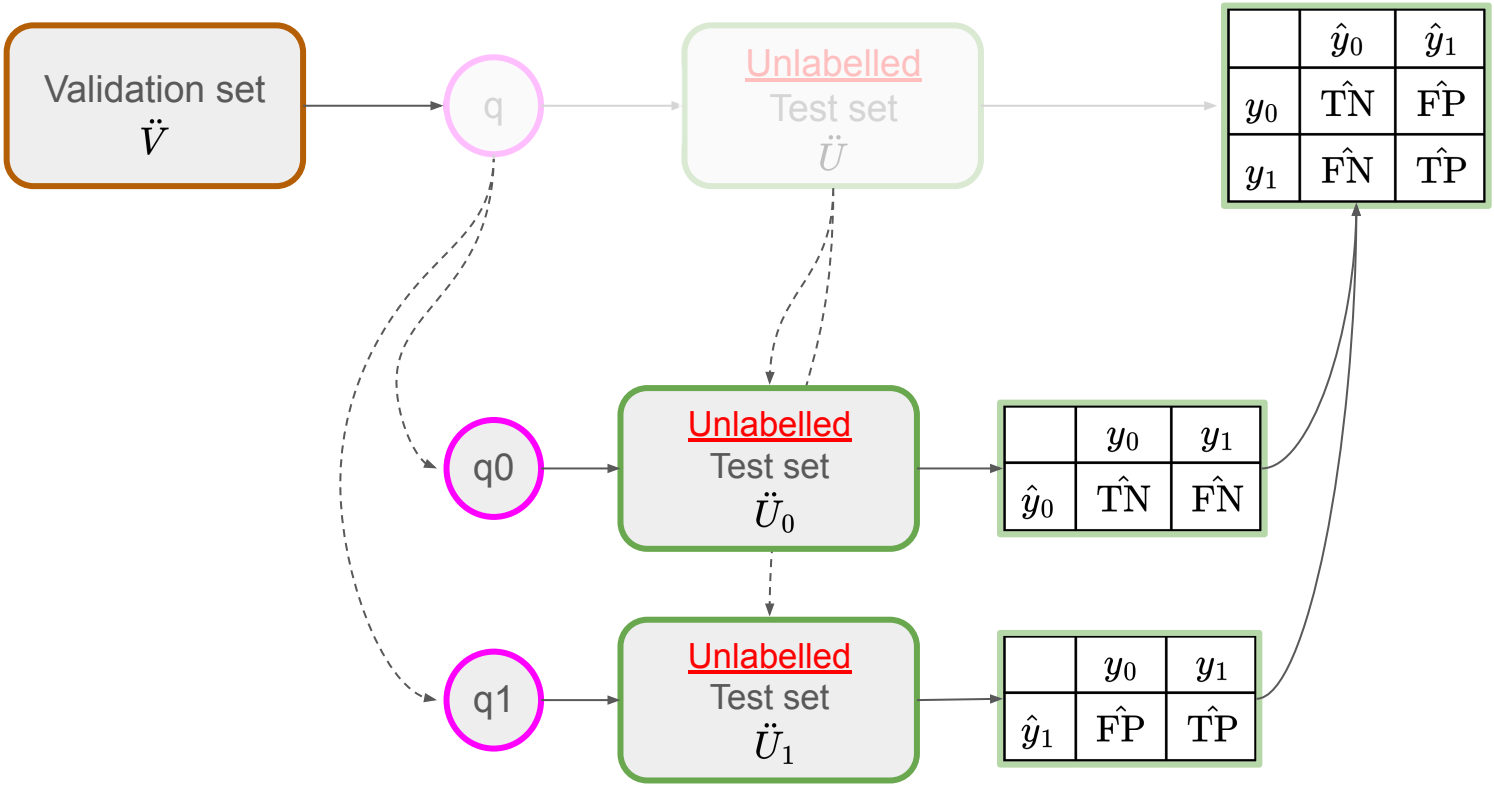
$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}}$$

$$F_1 = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}}$$

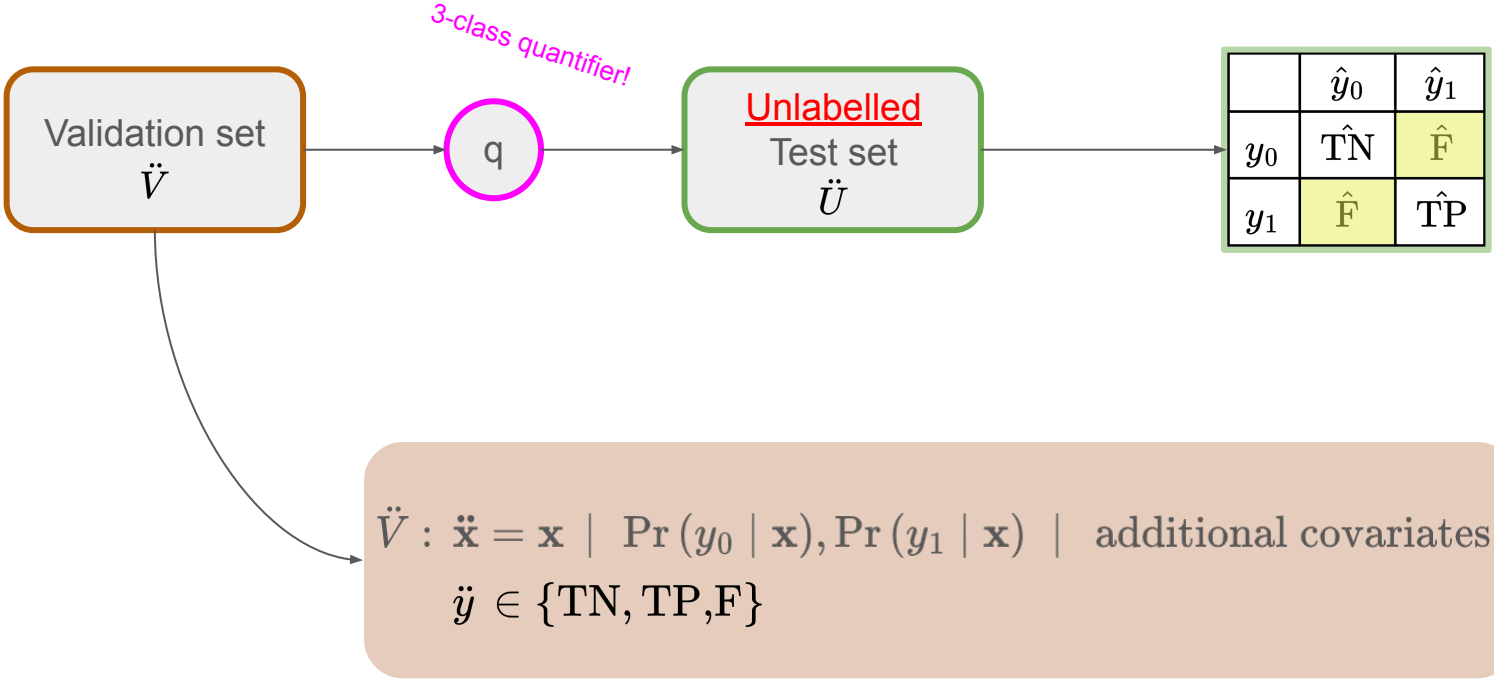
QuAcc - 1x4


Validation set
$\ddot{V}$
q
Unlabelled
Test set
$\ddot{U}$
$\hat{y}_0$
$\hat{y}_1$
y_0
$\hat{\text{TN}}$
$\hat{\text{FP}}$
y_1
$\hat{\text{FN}}$
$\hat{\text{TP}}$

QuAcc - 2x2


Validation set
$\ddot{V}$
q
Unlabelled
Test set
$\ddot{U}$
$\hat{y}_0$ $\hat{y}_1$
y_0 $\hat{\text{TN}}$ $\hat{\text{FP}}$
y_1 $\hat{\text{FN}}$ $\hat{\text{TP}}$
q0
Unlabelled
Test set
$\ddot{U}_0$
y_0 y_1
$\hat{y}_0$ $\hat{\text{TN}}$ $\hat{\text{FN}}$
q1
Unlabelled
Test set
$\ddot{U}_1$
y_0 y_1
$\hat{y}_1$ $\hat{\text{FP}}$ $\hat{\text{TP}}$

QuAcc - 1x3


3-class quantifier!
Validation set
$\ddot{V}$
q
Unlabelled
Test set
$\ddot{U}$
$\hat{y}_0$
$\hat{y}_1$
y_0
$\hat{\text{TN}}$
$\hat{\text{F}}$
y_1
$\hat{\text{F}}$
$\hat{\text{TP}}$
$\ddot{V} : \ddot{\mathbf{x}} = \mathbf{x} \mid \Pr(y_0 \mid \mathbf{x}), \Pr(y_1 \mid \mathbf{x}) \mid$ additional covariates
$\ddot{y} \in \{\text{TN}, \text{TP}, \text{F}\}$

Results

- http://ilona.isti.cnr.it:33421/plot