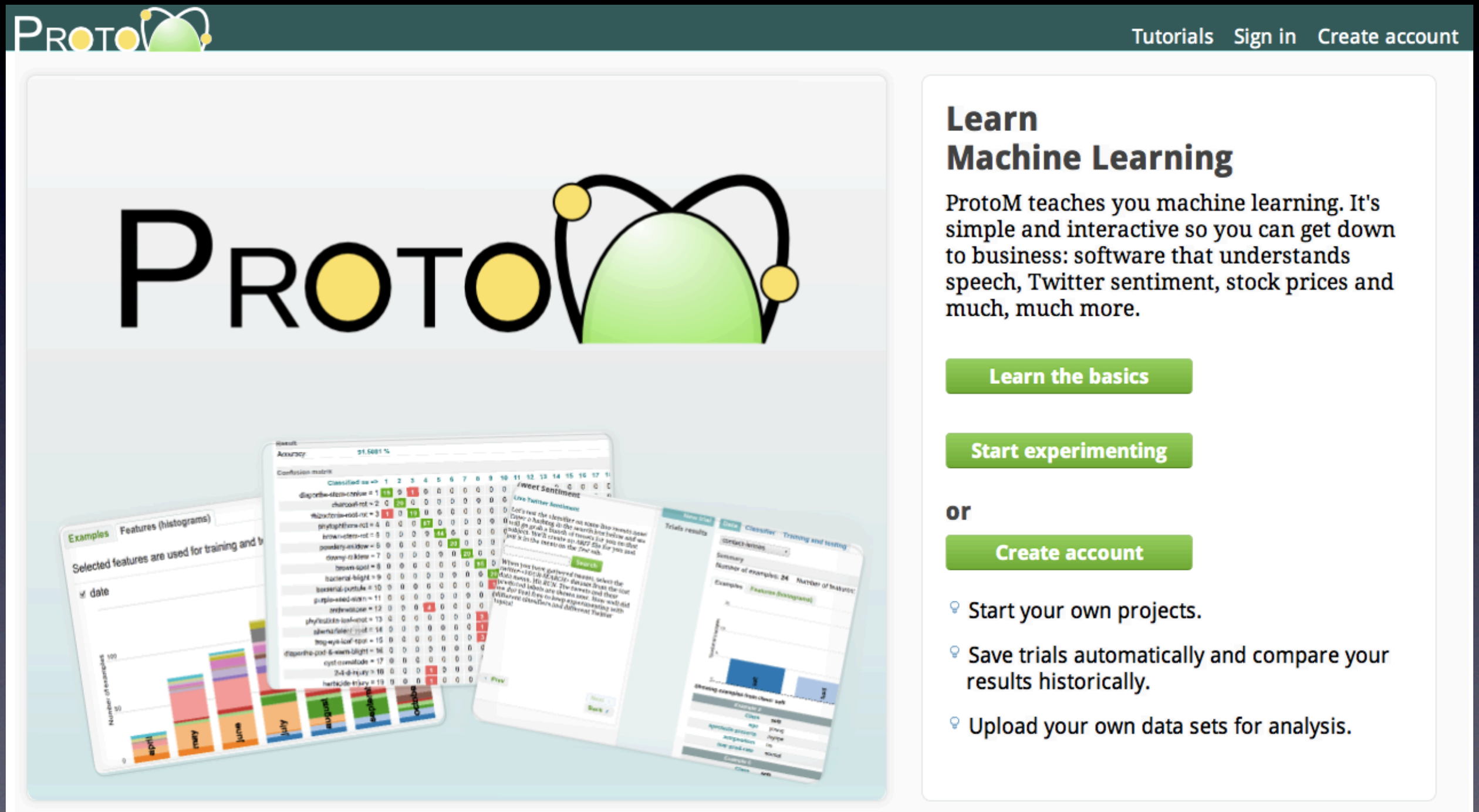
ProtoM teaches you machine learning



Through interactive tutorials

PROTO

Tutorials Sign in Create account

Tweet Sentiment

Choose a topic!

Let's test the classifier on some live tweets now! Enter a topic in the search box below and we will go grab a bunch of related tweets for you. The tweets will be saved as a new dataset in the *Use test data* menu.

Search

When you have gathered tweets, select the *twitter-<YOUR-SEARCH>* dataset from the *Use test data* menu. Hit *RUN*. The tweets and their predicted labels are shown now. How well did we do? Feel free to keep experimenting with different Twitter topics!

‹ Prev Next ›

Back

New trial

Trials results

× Trial-1

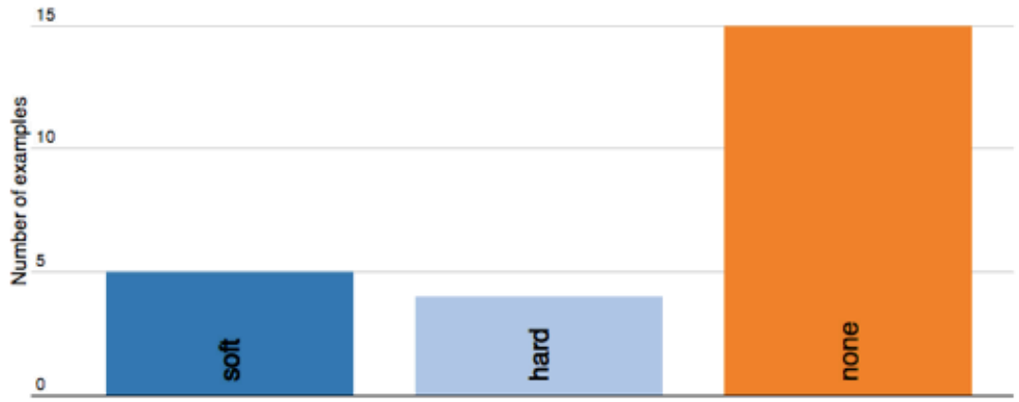
Data Classifier Training and testing

contact-lenses

Summary

Number of examples: **24** Number of features: **4** Class label: **contact-lenses**

Examples Features (histograms)



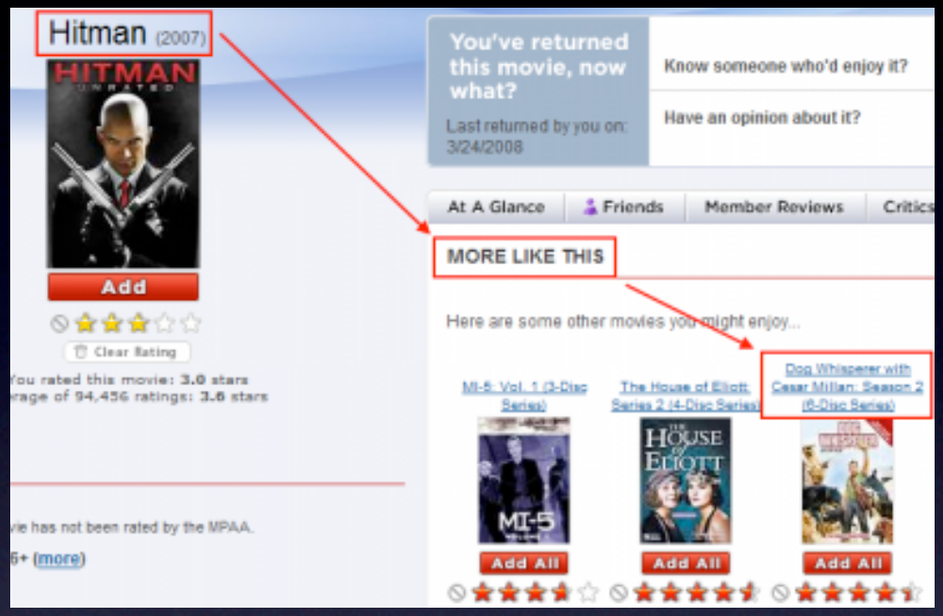
Showing examples from class: soft

Example 2	
Class	**soft**
age	young
spectacle-prescrip	myope
astigmatism	no
tear-prod-rate	normal

Example 6	
Class	**soft**
age	young
spectacle-prescrip	hypermetrope
astigmatism	no
tear-prod-rate	normal

Machine learning's everywhere

Speech recognition



Recommendation

Face recognition

Traditional machine learning course

Finding a decision hyperplane: the perceptron algorithm

Rosenblatt, 1957

Find the linear separator, characterized by parameters θ, that minimizes **training error**:

$$\hat{R}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\Big(1 - \delta(y_t, f(x_t;\theta))\Big) = \frac{1}{n}\sum_{t=1}^{n}\text{Loss}(y_t, f(x_t;\theta))$$

where $\delta(y,y') = 1$ if $y = y'$ and 0 otherwise.

For simplicity, we use the **zero-one loss** that is 1 for mistakes and 0 otherwise.

Perceptron update rule

- Start with any value for θ ($\mathbf{0}$ is typical)
- Go through training examples x^t, y^t one by one
- Update
  $\theta' \leftarrow \theta + y^t x^t$ if $y^t \neq f(x^t;\theta)$
- Stop when entire training set is categorized correctly

Abstract and theoretical

Whats happening?

Parameter updates tend to correct mistakes

When we make a mistake on x^t,

- the sign of $\theta \cdot x^t$ disagrees with y^t
- the product $y^t(\theta \cdot x^t)$ is negative
- the updated parameters are $\theta' = \theta + y^t x^t$

Now, if we try to classify x^t with the new parameters, we will have

$$\begin{aligned} y^t(\theta' \cdot x^t) &= y^t(\theta + y^t x^t) \cdot x^t \\ &= y^t(\theta \cdot x^t) + y^{t2}(x^t \cdot x^t) \\ &= y^t(\theta \cdot x^t) + |x^t|^2 \end{aligned}$$

The value of $y^t(\theta \cdot x^t)$ increases as a result of the update (becomes more positive). Successive updates based on x^t will change θ so that it is eventually classified correctly.

What happens when we cycle through the training examples?

Analysis of the perceptron algorithm

- Stops updating parameters when all examples are correctly classified
- Will always find a separator if one exists, in a finite number of updates
- Will not terminate if data are not separable
- Number of updates is related to the **margin**: how close the separating boundary is to the points

Margin of example i with separator θ^*:

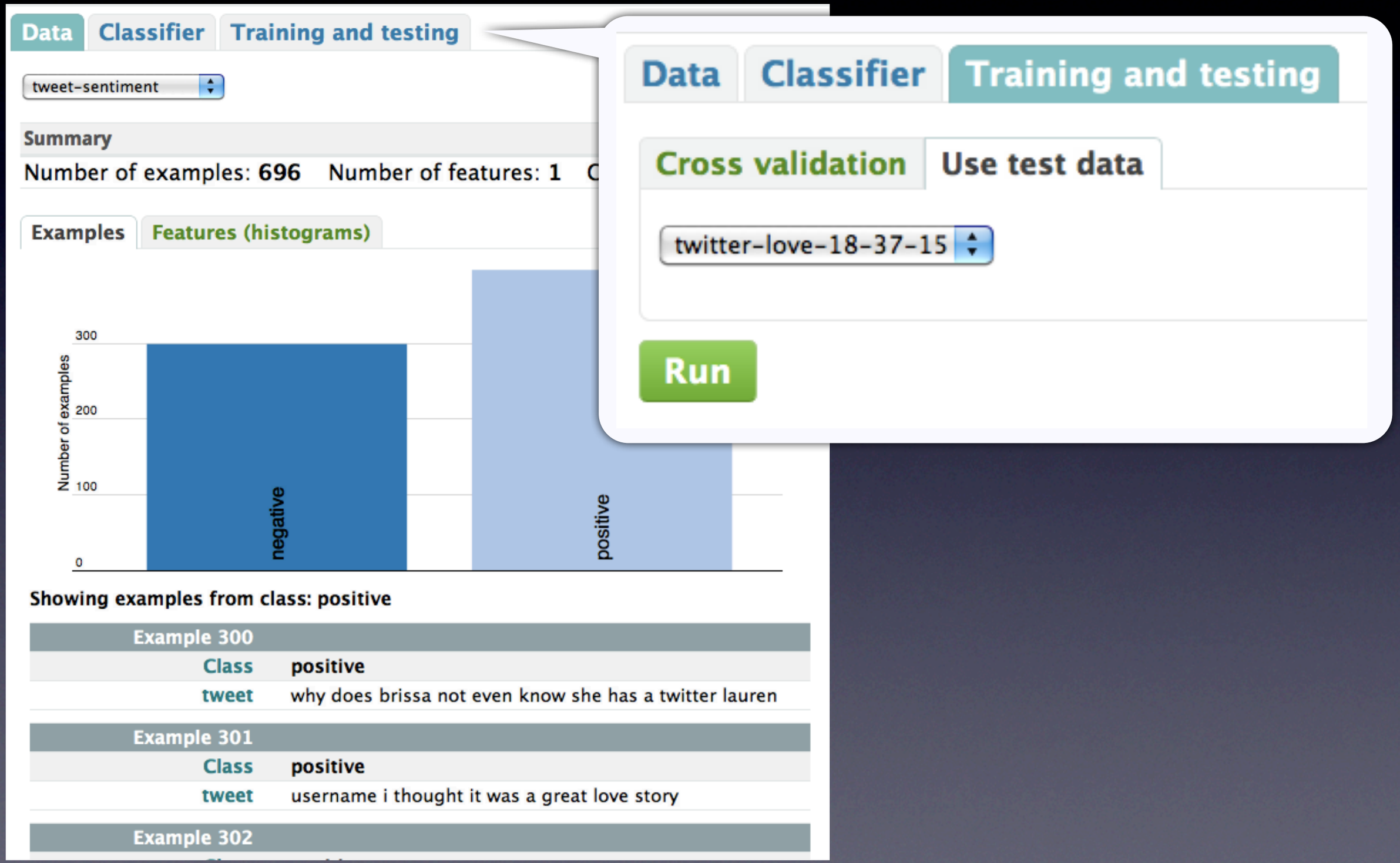
$$\frac{y^i(\theta^* \cdot x^i)}{||\theta^*||}$$

Margin of the dataset is the minimum of the margins of the examples

What is ML all about?
An example: Tweet sentiment

Training data

Practical ML with ProtoM



Practical ML with ProtoM

Predicted results

Predicted classification		
ID	prediction	tweet
1	positive	i love my girlfriend username she is the bestt
2	negative	had the volume up to hear username forgot to turn it down nearly burst my eardrums when i put find your love on that was silly of me
3	positive	username please follow me i love you so so much please i love you
4	positive	i i love you like a love song babyy
5	positive	username love the show any chance of a follow back
6	positive	username hahaha i love how much attitude you have here looks great

ProtoM makes ML accessible and fun

PROTOM

clevershadow@gmail.com Tutorials Projects Sign out

Machine Learning Basics

So, what is machine learning?

It's all about teaching computers to see patterns in data. We see patterns in data all the time. If I show you a picture of a blimp, you'll recognize it as a blimp because you've seen blimps in the past and you've learned that blimps are large, egg-shaped flying machines.

Well, computers have none of this intelligence built-in. Machine learning is all about training software to see data as blimps, cars, people, spoken words, stock prices, and almost anything else you can think of.

‹ Prev Next ›

Back

New trial

Trials results

× Trial-1

Trial result

Summary

Data:	**contact-lenses**	Classifier:	**rules.ZeroR**
Examples:	**24**	Class:	**contact-lenses**
Test mode:	**Cross validation**	Test data:	
Features used:	**4 / 4** age spectacle-prescrip astigmatism tear-prod-rate		

Result

Accuracy: 62.5000 %

Confusion matrix

Classified as =>	1	2	3
soft = 1	0	0	5
hard = 2	0	0	4
none = 3	0	0	15