
NETFLIX
Netflix Prize
Home | Rules | Leaderboard | Register | Update | Submit | Download
NETFLIX
Browse | Recommendations | Friends | Queue | Buy DVDs
Movies For You
You really liked it...
Welcome!
The Netflix Prize seeks to substantially improve the accuracy of predictions about how much someone is going to love a movie based on their movie preferences. Improve it enough and you win one (or more) Prizes. Winning the Netflix Prize improves our ability to connect people to the movies they love.
Read the Rules to see what is required to win the Prizes. If you are interested in joining the quest, you should register a team.
You should also read the frequently-asked questions about the Prize. And check out how various teams are doing on the Leaderboard.
Good luck and thanks for helping!
FAQ | Forum | Netflix Home
© 1997-2006 Netflix, Inc. All rights reserved.

Is there any customer information in the dataset that should be kept private?

No, all customer identifying information has been removed; all that remains are ratings and dates.

This follows our privacy policy, which you can review here.

Even if, for example, you knew all your own ratings and their dates you probably couldn't identify them reliably in the data because only a small sample was included (less than one-tenth of our complete dataset) and that data was subject to perturbation.

Of course, since you know all your own ratings that really isn't a privacy problem is it?

Terminator | Inside Out | Sound of Music | XXX | Michael Moore

Customers

12796 ~~fred~~ 4 0 0 1 2 0 ...

⋮

85632 ~~Iain~~ 5 ~~5~~4 ~~2~~1 0 0 0 5 1 5 4 ...

⋮

IMDB

Iain 9 10 3 0 0 0 0 ...

Probabilistic Model

Likelihood:

$$P(D \mid \underline{w}) = \prod_{n=1}^{N} \cancel{p(x^{(n)} \mid \underline{w})}\; p(y^{(n)} \mid \underline{x}^{(n)}, \underline{w})$$

Often "1" in supervised models
We know $\underline{x}^{(n)}$

↑
Data $\{\underline{x}^{(n)}, y^{(n)}\}$

For each
can minimize
−ve log likehood
(+ regularize?)

Examples

$\mathcal{N}(y^{(n)}; f(\underline{x}^{(n)}; \underline{w}), \sigma_y^2)$

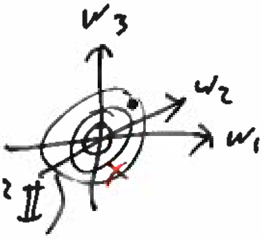
$\mathcal{N}(y^{(n)}; f(\underline{x}^{(n)}; \underline{w}), e^{g(\underline{x}^{(n)}; \underline{w})}$

$\text{Bernoulli}(y^{(n)}; \sigma(f(\underline{x}^{(n)}; \underline{w})))$

Robust version

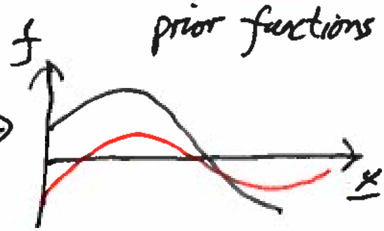
Prior

$p(\underline{w})$

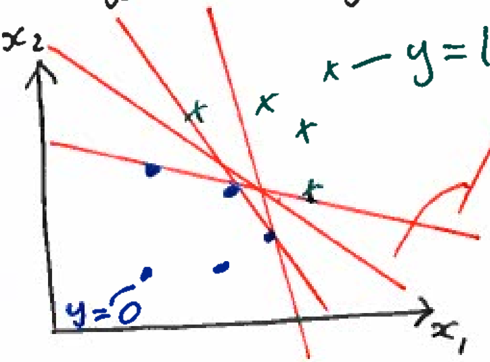


$\mathcal{N}(\underline{0}, \sigma_w^2 \mathbb{I})$

$\Rightarrow$



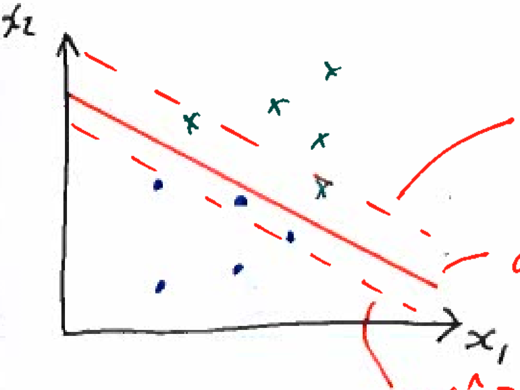
Bayesian Logistic Regression



Decision boundaries:

$p(y=1 \mid \underline{x}, \underline{w}) = \sigma(\underline{w}^T \underline{x}) = \frac{1}{2}$

for different plausible $\underline{w}$

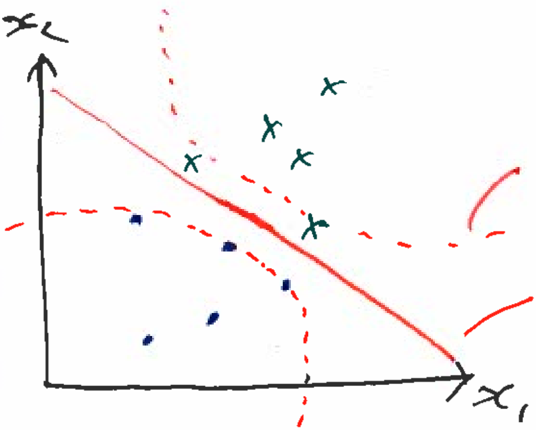


$\sigma(\hat{\underline{w}}^T \underline{x}) = 0.9$

$\sigma(\hat{\underline{w}}^T \underline{x}) = 0.5$

$\sigma(\hat{\underline{w}}^T \underline{x}) = 0.1$

Predictive distribution for a single fitted $\hat{\underline{w}}$, eg L2 regularised fit.



$p(y=1 \mid \underline{x}, D) = 0.9$

$p(y=1 \mid \underline{x}, D) = 1/2$

↑ Training data (D)

Plausible weights described by Posterior

$$p(\underline{w} \mid D, M) = \frac{p(D \mid \underline{w}, M)\, p(\underline{w} \mid M)}{p(D \mid M)}$$

D: Train $\{\underline{x}^{(n)}, y^{(n)}\}$

$p(D \mid M) = \int p(D \mid \underline{w})\, p(\underline{w})\, d\underline{w}$

M, model choices, hyperparameters, basis f^{n}'s (Often miss out)

Likelihood

$$P(D \mid \underline{w}) = \prod_n \cancel{p(\underline{x}^{(n)} \mid \underline{w})}\; p(y^{(n)} \mid \underline{w}, \underline{x}^{(n)})$$

(1: "known" here)

$p(y^{(n)} \mid \underline{w}, \underline{x}^{(n)}) = \sigma(\underline{w}^T \underline{x}^{(n)} (2y^{(n)} - 1))$

Prior:

$$p(\underline{w}) = \mathcal{N}(\underline{w}; \underline{0}, \sigma_w^2 \mathbb{I})$$

Marginal Likelihood

$p(D)$ or $p(D \mid M)$ can compare models

Predictions

$$p(y=1 \mid \underline{x}, D) = \int p(y=1, \underline{w} \mid \underline{x}, D)\, d\underline{w}$$

$y=1$: test output; $\underline{x}$: test input

$$= \int p(y=1 \mid \underline{w}, \underline{x}, \cancel{D})\, p(\underline{w} \mid \cancel{\underline{x}}, D)\, d\underline{w}$$

$p(y=1 \mid \underline{w}, \underline{x}) = \sigma(\underline{w}^T \underline{x})$; $p(\underline{w} \mid D)$: posterior