# **Overfitting in classification**

## **Evaluating a classifier**

Diagram

Description automatically generated

Sample-1 from validation data  
Diagram

Description automatically generated  
Initially, give Sushi was great and it has +ve review.

Timeline

Description automatically generated with medium confidence  
I feed the validation sample into the classifier.  
It predicts the output. (y^=positive).  
The validation sample has the positive output. [Both are matched]  
So correct=1.

Sample-2  
Graphical user interface

Description automatically generated with medium confidenceA picture containing diagram

Description automatically generated

I feed the validation sample into the classifier.  
It predicts the output. (y^=positive).  
The validation sample has the negative output. [Both are mis-matched]  
So mistakes=1.

Text

Description automatically generated

## **Review of over-fitting in regression**

Chart, line chart

Description automatically generatedDiagram

Description automatically generated with medium confidence

## **Over-fitting in classification**

Scatter chart

Description automatically generated

Lets try some quadratic fit

Diagram

Description automatically generated

This seems to be good, but lets try some higher degree polynomial.

## **Visualizing over-fitting with high degree polynomial features**

Diagram

Description automatically generated  
The error over here is 0.  
The coefficients are extremely large in terms of magnitude.  
These are the early sign of over-fitting.

A picture containing diagram

Description automatically generated

Diagram

Description automatically generated

# **Over confident predictions due to overfitting**

## **Overfitting in classifiers leads to overconfident predictions**

Diagram

Description automatically generated

A picture containing text

Description automatically generated

As the model becomes extremely over-fit, the model will predict the output as exactly 1 or 0 without having any doubt [more over-confident].

Chart

Description automatically generated with low confidence

Keeping the same threshold=1  
Probability of a review to be +ve 🡪 0.73  
Probability of a review to be +ve 🡪 0.88  
Probability of a review to be +ve 🡪 0.997

As the coefficients are getting bigger and bigger every time, the estimated probability of the review becomes more-more steeper.

Not only the curve looks very weird but the estimated probability becomes close to 0 or close to 1.

So lets look our data-set and see how we observe the same effect.

## **Visualizing over-confident predictions**

Chart

Description automatically generated with medium confidence

* This white region is where the probability is 0.
* The entire diagram is the plot of P(y=+1).
* The points to the top left P(y=+1) ~ 0.
* The points to the bottom right P(y=+1) ~ 1.
* Wide Uncertainty region, the P(y=+1) ~ 0.5
* Though the linear classifier is not a great fit to the data (cannot able to classify the points as +ve/-ve), the uncertainty region makes quite a lot of sense.
* In the narrow white region, the points are mis-classified.
* So I am highly uncertain whether the points are +ve/-ve.

A picture containing diagram

Description automatically generated

* The entire diagram is the plot of P(y=+1).
* The quadratic fit seems to be a better fit to the data.
* This is really a great fit not in-terms of decision boundary but also in-terms of probability.
* The places where the probabilities are close to 0.5 are really the ones where we are unsure what is going on.
* We don’t know whether the points are -ve / +ve classified on the left / righ side of the parabola.

Diagram

Description automatically generated

The uncertainty region is extremely thin. So we should be extremely sure to make a point as +ve/-ve [over-confident]

Uncertainty is really important in classifiers and can be avoided by creating a narrow bands instead of a thin one.

## **Another perspective on overfitting in logistic regression** Chart, scatter chart Description automatically generatedDiagram Description automatically generatedA picture containing diagram Description automatically generatedA picture containing text, document Description automatically generatedText Description automatically generated with medium confidenceText, letter Description automatically generatedText, letter Description automatically generatedDiagram Description automatically generated

# **L2 regularized Logistic Regression**

## **Penalizing the large coefficients to mitigate over-fitting**

Diagram

Description automatically generated

A picture containing text

Description automatically generatedDiagram

Description automatically generatedText

Description automatically generated

Timeline

Description automatically generated

## **L2 regularized Logistic Regression**

Text

Description automatically generated with low confidence

Text

Description automatically generated

Graphical user interface, text, application

Description automatically generated

## **Visualizing effect of regularization on logistic regression**

Chart, scatter chart

Description automatically generatedGraphical user interface

Description automatically generated with medium confidence

Chart

Description automatically generated



Graphical user interface

Description automatically generated



## **Learning L2 regularized logistic regression with gradient ascent**

Diagram

Description automatically generated

Diagram

Description automatically generated with medium confidence

Diagram

Description automatically generated with low confidence

Text

Description automatically generated

# **Sparse Logistic Regression**

# **Summarizing overfitting and regularization in logistic regression**

