* Trouble shooting *

Assume that your hypothesis not performing well on the test set.

But how will i know WELL that it's not performing?

1. Classify your training data

train set $\underbrace{\quad}_{\approx 70\%}$ Test set. $\underbrace{\quad}_{30\%}$

Evaluate error on this using hypothesis generated from training set.

For linear R use squared error

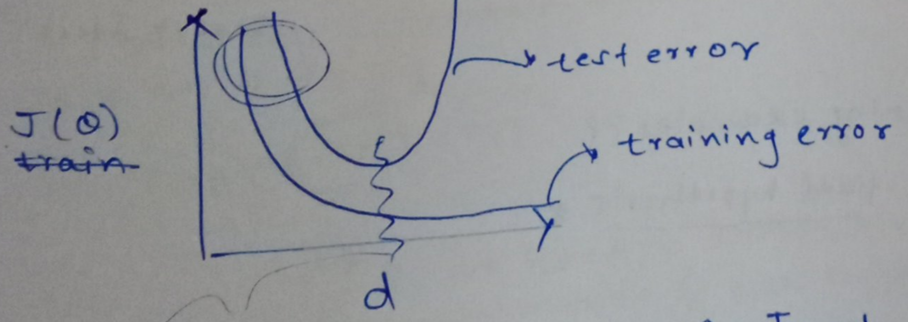
For classification use accuracy $\frac{\text{correct pred}}{\text{total pred}}$.

Zalqlav: How to choose the best order model. {cross validation not introduced yet}

A.
1. $\Theta_0 + \Theta_1 x$ $d = 1 \rightarrow$ Train on this set $\rightarrow \Theta^{(1)}$
2. $\Theta_0 + \Theta_2 x + \Theta_2 x^2$ $d = 2 \rightarrow$ - - - $\rightarrow \Theta^{(2)}$
3. ⋮ $\Theta^{(3)}$

$\Theta^{(4)}$

B. Now use the Θs obtained before;



When the

$J_{train} \approx J_{test}$

$\Rightarrow$ Probably {Bias}

$J_{test} \gg J_{train}$

$\Rightarrow$ Probable {Variance} overfitting

CHOOSE ME BUT DONT REPORT ME

Because, "I am very very optimistic!!" { Report me on test set.

→ set me on cross validation set.

(C) How to report regularisation?

Note that you use gradient descent on model

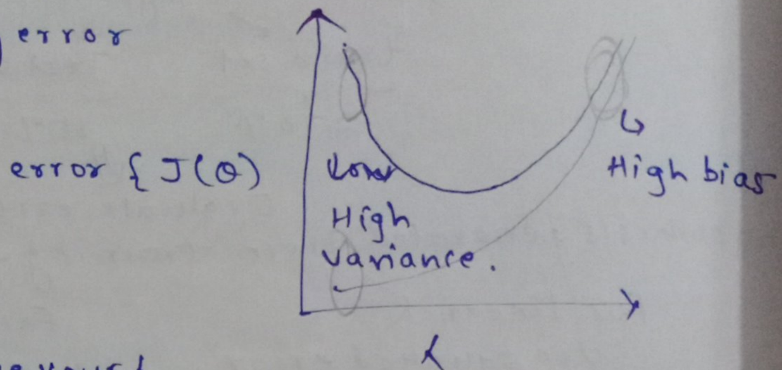
$$J(\theta) = \text{Sqerror} + \lambda_1 \theta^2 \rightarrow \theta^{(t)}$$

$\downarrow$ train

But report cost in terms

$J''_{train}(\theta) = \text{Sqerror}$

$J''_{cv}(\theta) = \text{sqerror}$



$J''_{train}(\theta) \approx J''_{CV}(\theta)$ (~~Test~~ CV)

Decrease your λ

$J''_{cv}(\theta) \gg J''_{train}(\theta)$ } Increase your λ

Learn θ_s through $J(\theta) = sq^2 + \lambda\theta^2 \in \lambda = \{0.01 \ldots 10.0\}$

Then using the θ_s obtained, plot

$J_{cv}(\theta) = sq^2 + \cancel{\lambda\theta^2}$ and obtain the graph as above.

---

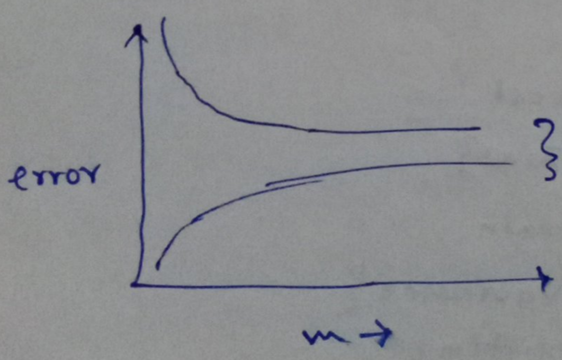
* The best combo *

Finding minimum through $d \Rightarrow$ Finding minimum through λ. } best combo of θ and λ

Report $J_{test}(\theta)$

(D) How about training examples??

* For a fixed hypothesis *



} If not yet converged

↓

Check for high variance

If converged

↓

No point in increasing training examples.

Rulebook.

1. Getting more training examples : Fixes high variance
2. Trying smaller set of features : Fixes high variance
3. Adding features : fixes high bias
4. Adding polynomial features : Fixes high bias
5. Decreasing λ : Fixes high variance
6. Increasing λ : Fixes high bias

---

* Getting real deep into System Design. *

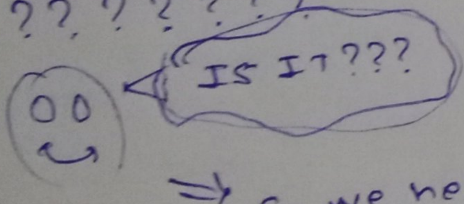
1: Getting the features to be selected right.
Use your own brain.

2: Protyping and analysing.
1) Start with a simple algorithm, implement it, Quick and dirty test it.
2) See if what is the problem, high bias, variance, training examples
3) Manually examine also.

Skewed Classes { Keyword : Skwed }
{ Dected

* Precision and Recall *

let us assume a cancer classification problem.
My hypothesis gives 99% accuracy
And actual cases of cancer are 0.5%.
And change my hypothesis to:
So i innovate and change my hypothesis to:
$y = 0$.
Now my accuracy is 99.5% { Great improvement
right ?? ? ? ? ? ?



$\Rightarrow$ So we need a better parameter. to equalate the performance of our hypothesis.
Specially important in case of skewed cases

		Actual		
		Positives	Negative	Precision
Predicted	Positive	True positive	False positive	$\frac{\text{T act. True}}{\text{Total pred True}}$
	Negative	False Negative	True Negative	
	Recall ⇒	$\frac{\text{Pred true}}{\text{Actual True.}}$		

∴ Best F value $2 * \frac{RP}{R+P}$ wins } Note: Twitch with threshold in case of LR and see the effect