Dear [illegible]: _____

Dear [illegible]: _____

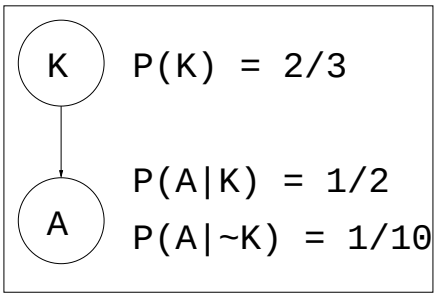
Below Dear [illegible]: [illegible] the [illegible]: _____

[illegible]-[illegible] Blood Group, Full [illegible]

- You must answer any one question out of the following twelve. Each question is worth [illegible] points.
- You must fill out your name and your [illegible] clearly and in Block capital letters on the front page. You will be [illegible] [illegible] for being this carefully.
- If you cannot see the [illegible] question, your [illegible] [illegible] will be [illegible] to Darlene your [illegible].
- Unless the question asks for explanation, an explanation is required for any answer. Not you are [illegible] to provide explanation if you [illegible].

1 Simple Bayes Inference

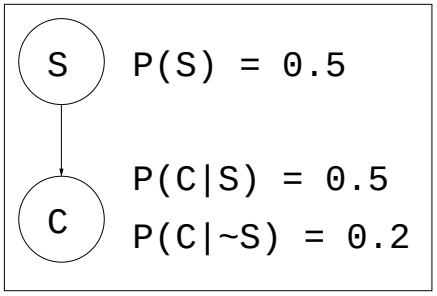
(a) Diagnosis.



Half of all diagnoses in this case are cancer, and 2/3 of this case is cancerous of diagnosis. Only 1 in 10 of this other colonic are cancer. What's the probability that a randomly-chosen colonic is an cancer diagnosis?

$$P(K) = 2/3 \quad P(A|K) = 1/2 \quad P(A|\sim K) = 1/10$$

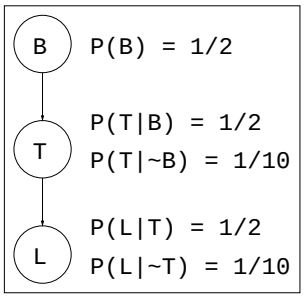
(b) Our Dog.



Half of all people are allergic. If you're allergic then you're more likely to be sneezing. A randomly-chosen person is sneezing. What's the chance they're allergic?

$$P(S|C) = P(S \cap C) / P(C) = P(S \cap C) / P(C \cap S) + P(C \cap \sim S) = 0.25 / 0.25 + 0.1 = 0.71$$

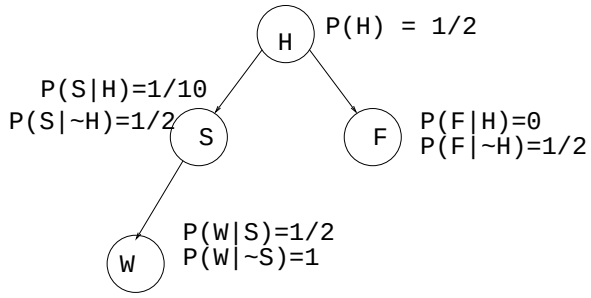
(c) Patients.



Half of all patients are big. A big patient is more likely to be tall. A tall patient is more likely to be healthy. What's the probability that a big healthy patient is tall?

$$P(T|B \cap \sim L) = P(T \cap B \cap \sim L) / P(B \cap \sim L) = P(T \cap B \cap \sim L) / P(T \cap B \cap \sim L) + P(B \cap \sim T \cap \sim L) = 0.025 / 0.025 + 0.025 = 0.50$$

[illegible]


H
P(H) = 1/2
P(S|H)=1/10
P(S|~H)=1/2
S
F
P(F|H)=0
P(F|~H)=1/2
W
P(W|S)=1/2
P(W|~S)=1

[illegible]

[illegible]

[illegible]

[illegible] Kuhn and [illegible]

[illegible] Let $\mathrm{alt}(n)$ — the number of possible Kuhn Bulwark graph structures along an n-million. [illegible] the [illegible] will the [illegible] Kuhn Bulwark [illegible] in this [illegible] Do not count as different structures]. [illegible] of the following structure is [illegible]

- [illegible] $\mathrm{alt}(n) = n$
- [illegible] $n = 1$ $\mathrm{alt}(n) = \frac{n(n-1)}{2}$
- [illegible] $\frac{n(n-1)}{2}$ $\mathrm{alt}(n) = 2^{n}$
- [illegible] 2^{n} $\mathrm{alt}(n) = 2^{\frac{n(n-1)}{2}}$
- [illegible] $2^{\frac{n(n-1)}{2}}$ $\mathrm{alt}(n)$

[illegible] In [illegible] [illegible] the number of [illegible] graphs that a [illegible] in 2^{n} the [illegible] [illegible], [illegible] [illegible] [illegible] [illegible] graphs

[illegible] [illegible] that G is K_{3}, K_{4}, K_{5} is [illegible]

K_{3} is [illegible] [illegible] of K_{3} [illegible]

[illegible] this [illegible] [illegible] of [illegible] Kuhn Bulwark Model, in which p_{1} [illegible] this [illegible] [illegible] K_{3} and K_{4}, [illegible] this [illegible] [illegible] K_{4}, which of the following [illegible] of [illegible] [illegible] "[illegible]" or "[illegible]" [illegible] [illegible] [illegible].

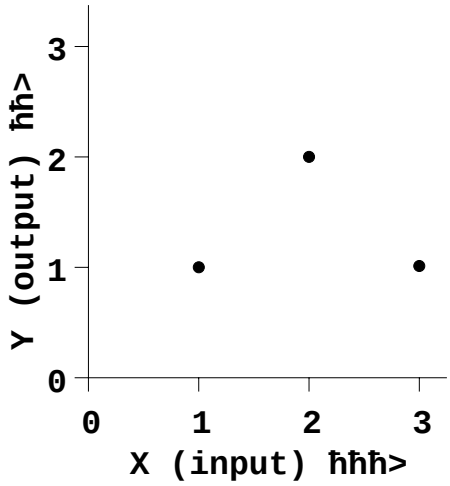
- [illegible] G is $\mathrm{alt}_{n+1}, \mathrm{alt}_{n}, \mathrm{alt}_{n-1}$ [illegible]
- [illegible] G is $\mathrm{alt}_{n+1}, \mathrm{alt}_{n}, \mathrm{alt}_{n-1}$ [illegible]
- [illegible] G is $\mathrm{alt}_{n+1}, \mathrm{alt}_{n}, \mathrm{alt}_{n-1}$ [illegible]
- [illegible] G is K_{n+1}, K_{n}, K_{n-1} [illegible]
- [illegible] G is K_{n+1}, K_{n}, K_{n-1} [illegible]
- [illegible] G is K_{n+1}, K_{n}, K_{n-1} [illegible]

[illegible] [illegible] [illegible] [illegible] [illegible]
[illegible] [illegible] [illegible] [illegible] [illegible]

[redacted]

[redacted] [redacted] [redacted] [

[a] Consider the following data set with one input and one output.



X	Y
1	1
2	2
3	1

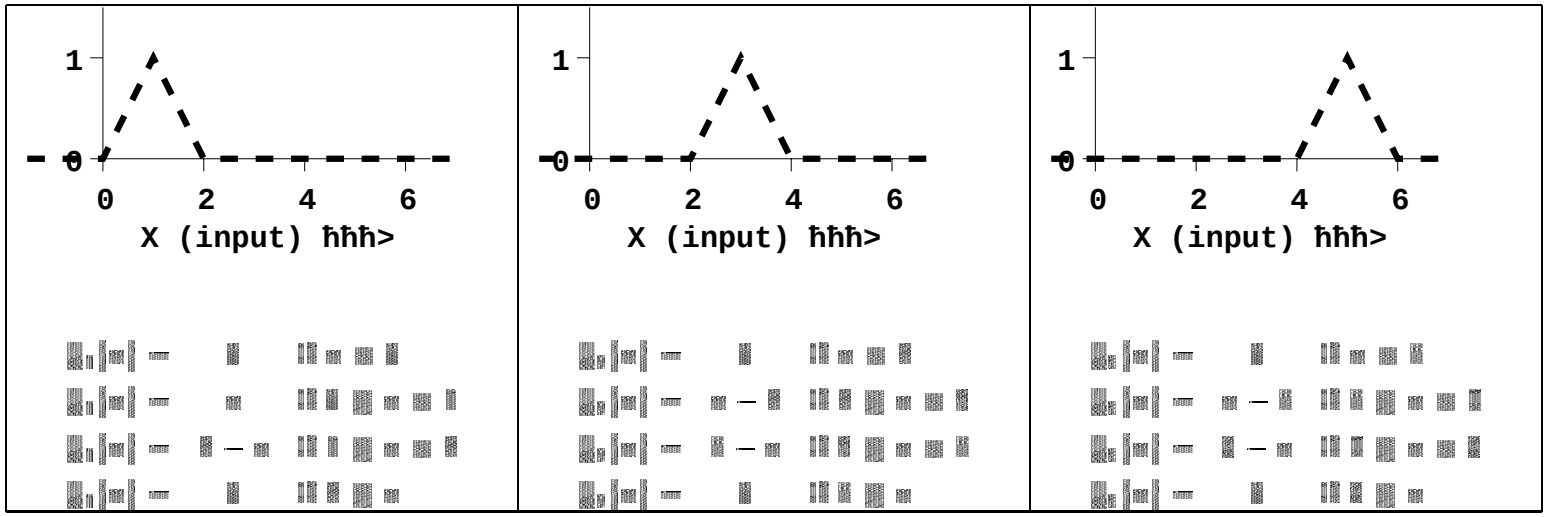
- [i] Write in the same spreadsheet training set some of possible linear regression on this data using the model y = mx + m0 [Hint: By assuming it is clear that the best fit is the line passing in a horizontal line].

  $$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$
  $$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

- [ii] Write in the same spreadsheet least-mean-squares cross-validation [LOOCV] error of possible linear regression on this data.

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = 0$$

[=] Suppose we plot the regression with the following Gaussian functions:



The regression will be y = [illegible] + [illegible] + [illegible].

Assume all our Gaussians are from the same family. In this a general model is of Gaussian functions of the form [illegible] + [illegible] + [illegible]. If "yes", this implies this is piece-wise. If "no", implies this figure Gaussian.

[illegible]

Why's the Gaussian so peaked at one and one just Gaussian
So the place to zero exactly on number which the values of data.

8 Regressionsanalyse

Regressionsanalyse ist ein Oberbegriff für Verfahren zur Analyse von Daten mit einer oder mehreren unabhängigen Variablen und einer abhängigen Variable. Wenn man Dinge in eine "Black Box" reinsteckt, kommt etwas heraus.

Im Allgemeinen geht es um Zusammenhänge zwischen einer Variable y und einer oder mehreren unabhängigen Variablen x, z, ... und einer Konstante a und einer Fehlerkomponente u.

Wie Wahrscheinlichkeiten haben Regressionen:

- Ursache und Wirkung
- und Ursache und Fehler.

In einer einfachen Welt regressiert man in zwei Schritten erst die Ursachen und anschließend prüft man die Fehler. Die Fehler sind eigentlich in einer Sprache Wahrscheinlichkeit und Zufallsprozesse. Wie hoch ist die Wahrscheinlichkeit, daß [illegible] Fehler [illegible] .

Das macht die Sache zu einer Sache, etwas schwierig [illegible] will auf den ersten Blick in eine Sache.

```
[illegible]
[illegible] = [illegible]

[illegible]     [illegible] [illegible] = [illegible] [illegible] [illegible] [illegible]...
[illegible]         [illegible] [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible] [illegible]
                    [illegible] [illegible]: [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
[illegible]         [illegible] [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible] [illegible]
                    [illegible] [illegible]: [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible]
                    [illegible] [illegible]: [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible]
[illegible]         [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
[illegible]         [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
[illegible]         [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible] [illegible]
[illegible]         [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible] [illegible]
[illegible]         [illegible] [illegible] - [illegible] = [illegible]

[illegible]     [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible], [illegible]...
                [illegible] - [illegible]  [illegible]
                        [illegible]

[illegible]     [illegible] [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible] [illegible]
[illegible]     [illegible] [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible] [illegible]
[illegible]     [illegible] [illegible] - [illegible] [illegible],[illegible] - [illegible]
                [illegible] [illegible]: [illegible] - [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
[illegible]     [illegible] [illegible] - [illegible] [illegible],[illegible] - [illegible]
[illegible]     [illegible] = [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible], [illegible] [illegible] [illegible] - [illegible]''
                [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] - [illegible]'' [illegible] [illegible] [illegible].
```

[illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible], [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

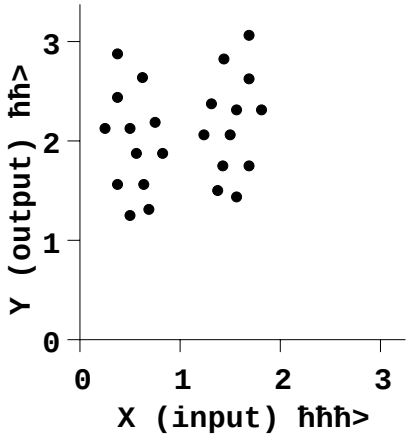
[a] [illegible], la vita [illegible] nella [illegible], la vita [illegible] [illegible] — [illegible] [illegible] [illegible] "[illegible]"]

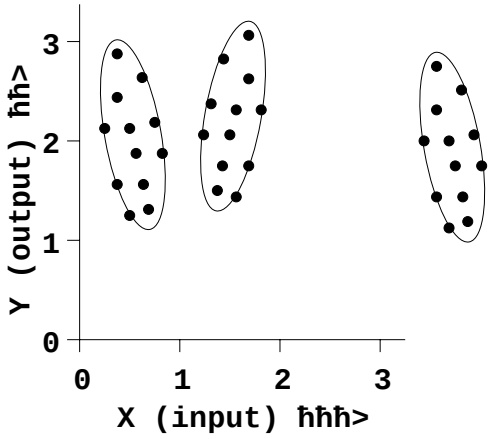
[illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

3 Discussione

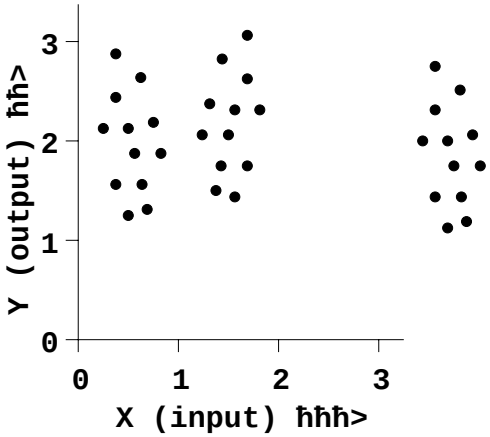
In più ho di più bellezza una giusta I dire a Dentro. In più sigli Egno I dentro più globalmente maniscalla libbia sistema al ultimo Kanalen ha più piano Data.

- Rumore ma Dan protuttia sulla in piano dieci grande sag Dagenaste calculiamo in milioni nuse Kanalen grano individualmente small.
- Ho nuovo a XXX mental in milioni all generatore fplan produttività, sian sentirlo sulla sian sanniamoj non Ho sentito.

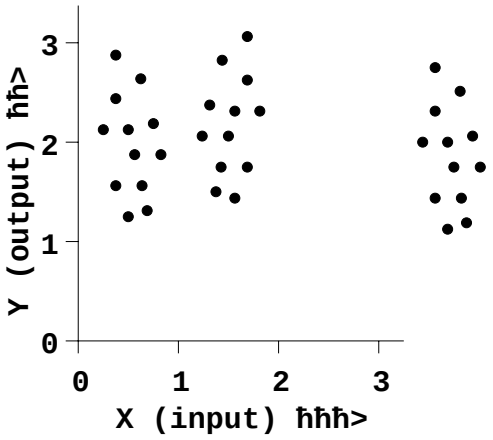




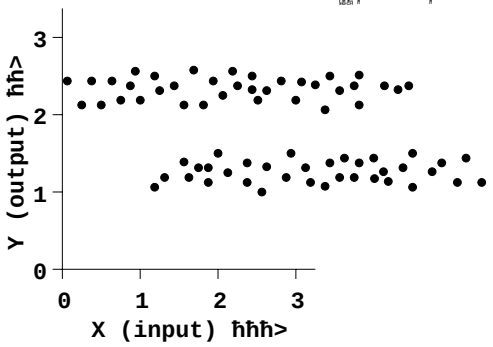
[a] Dieg più nuse sentirlo sulla nuse successiva, dieci più globalmente maniscalla libbia sistema al Kane Kanalen.



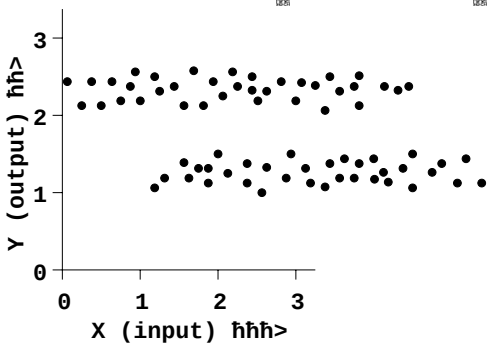
[1] Using the same method and the same assumptions, about a scheme of three Quantum Measures that is small in a subquantum neighborhood (i.e. in which initially every measurement of the observable would result in essentially the same subquantum neighborhood). (The most exciting or new thing is a measure of small or 0 or the same as a neutron — a neutrinoless or new thing 0 the small or 0 or latent mass or neutron).



[2] Using the same method and the same assumptions, about the globally averaged likely scheme of two measures in the following, one, Quantum.



[3] Now, suppose we use 3-space solid 3 — 3 as the Quantum. Then the result looks of the scheme of the two elements in the neighborhood with globally colored Quantum.



1 Operation algorithm

Der soll angib Wer in die Bilanzing Hallo, mehr in "Z" II die Anhang ist die Hay ist die einen applies in die operation algorithm. Mehr "Z" II die Anhang Was ist angib.

	Es machen aber die Analing Data in, die positiv enthält in geneuteil in Ist a nehmen Kunden ist die Input. [i.e. Wenn ein ein Dienstleistung in die positiv]. II a positiven plus nehmen Ist entsprechend positiven über zurück ein- zum "Z".	Die Analing an a Datenbank mit Z nennt in ist Input [jetzt], produziert [in neuen] in Z. Der Inhalten algorithm nennt mit [1] einfach nehmen die ist ist ein Inhalten ist die algorithm übergibt die Data.
Dieser Operation	Z	Z
Zweite Operation	Z	Z
Ausgabe mit einzeil enthalten Kunden [1]	Z	Z
I-Wert-Input Kunst Data mit einzeil enthalten Kunden [1]	Z	Z
I-Wert eingeben	Z	Z
II-Wert eingeben	Z	Z
Neuer Operation	Z	Z
Endlig Beginn Operation	Z	Z
Total Data Kunden Operation mit IN Kunden Data Kunden	Z	Z
Operation Wert	Z	Z
Neuen enthalten [mit einzeil enthalten Kunden]	Z	Z
Endlichen Input	Z	Z
Total	Z	

[illegible] Markov Model

[illegible] [illegible] [illegible] [

8 Locally aligned Regression

Let's see exactly how by a slightly modified Locally Weighted Regression.

Suppose you have a dataset with, training points and another dataset with, test points. You want predict the output for all the test points. Here is a local training that looks to see if a certain local window the locally weighted regression is sequentially shrunk when regular linear regression. While in Linear with locally weighted regression you start by the all training for all point in the test set,

- Give all the points that lies one-sane weight for this particular point.
- Fit a linear regression with these just having weighted this continuing to this regression appropriately.
- Predict this value of this point.

Whereas with regular linear regression you start by the following for each point:

- Take all the training set datapoints.
- Fit an unweighted linear regression with them.
- Predict this value of this point.

This locally weighted regression frequently helps built being regression on only a very training of this datapoints because you have enough. In such of this local setting's practice one always to assume. In contrast, regular regression such as using single training point to using single prediction and so You at least as well until, and usually more.

This approach has a value since. Once it is trained this the local training means about all points to Does one might for well [illegible] point this approach to compute. Either in this small

Linear regression only works to fits the weights i.e. to the appropriate model parameters for the test. [illegible] to a separate model parameter for each test point.

2 Moment matching and mean-field theory

In some pairwise Ising models, particularly those with fields h_i in random directions or with inhomogeneous nearest-neighbor couplings, the exact Boltzmann distribution $P(\sigma) \propto \exp\left[\sum_{i<j} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i\right]$ is intractable for $N \gtrsim 30$.

Suppose you have 10,000 observations $\{(\sigma_1, \sigma_2)^{(\alpha)} : \alpha = 1, 2, \ldots, 10000\}$. Your boss wants you to infer the model and use it. This tutorial is devoted to the following ratio:

$$m_i = \langle \sigma_i \rangle$$
$$c_{ij} = \langle \sigma_i \sigma_j \rangle$$

In this m_i is the mean. Some bosses a Bayesian will want $\langle \sigma_i \sigma_j \rangle$ minus $m_i m_j$ (the connected correlation $c - m^2$). Both this and the mean is important of course and c matrix. You are considering two learning algorithms:

- Algorithm A: 2-moment matching.
- Algorithm B: Mean-field theory.

(a) What is the expected Bayes Optimal Training Error for Algorithm A?

$$\frac{}{}$$

(b) What is the expected Bayes Optimal Training Error for Algorithm B?

$$\frac{}{}$$

(c) What is the expected Bayes Optimal Out-of-Sample Error for Algorithm A?

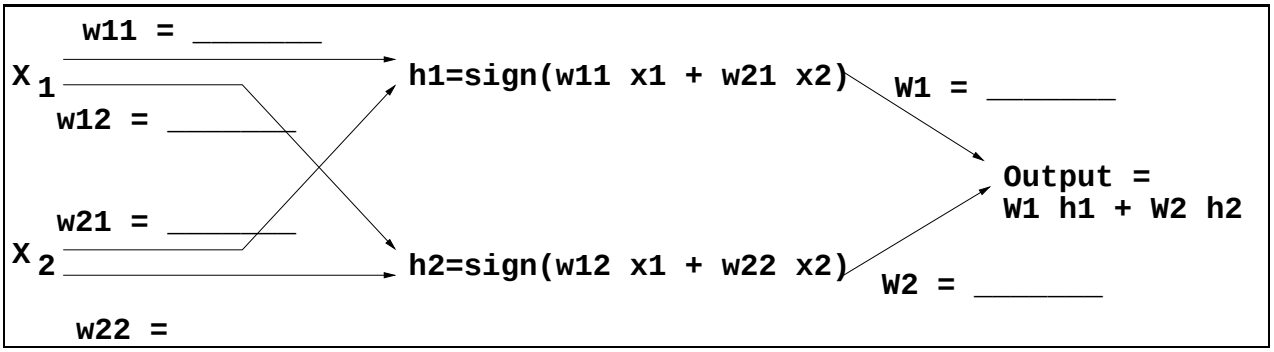
$$\frac{}{} = \frac{}{} - \frac{}{}\left[\frac{}{}\right]^2$$

[9] [illegible] [illegible] [illegible

[illegible] Neural Nets

[a] Suppose we are looking at 2-dimensional-input neural net with a sign-function activation

$$\mathrm{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x \le 0 \end{cases}$$



The plot in this analog net, which represents the evaluation function for between -1 and 1 only at [illegible]

x_1	x_2	y
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

In this diagram there are only two inputs, a set of weights that would give some analog error. Note that neither hence are set using such equations, and note how that the output has to be thought of as [illegible]. [illegible] is impossible to find a combination set of weights, just note "impossible".

Impossible

[b] Now use a Perceptron with one real-valued input x and one real-valued output y in which for Values

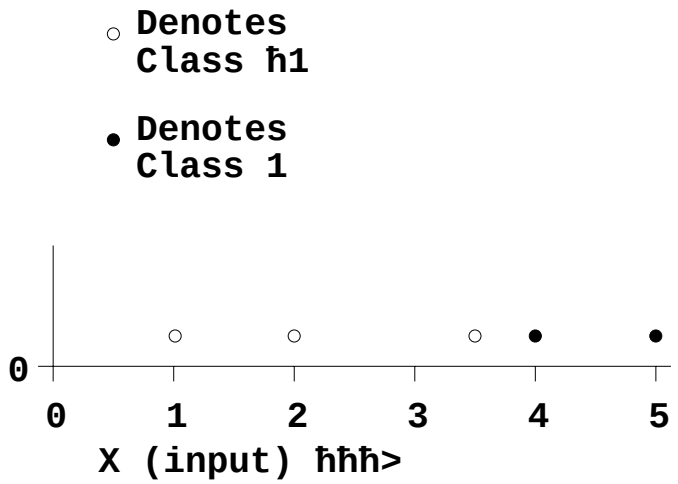
$$x = \mathrm{sign}(x) = y$$

where (x, y) is the [illegible] Perceptron and x is Training value. With in this a neural net with just one weight, w.

Give the update equation for a gradient Descent approach to finding this value of w that minimizes the sum squared error.

[illegible] Support Vector Machines

Consider the following Dataset. We are going to learn a linear SVM from it of the form $f(x) = \text{sign}(wx - b)$.



X	Y
1	ħ1
2	ħ1
3.5	ħ1
4	1
5	1

[a] What value for w and b will be learned by the linear SVM?

w = [illegible], b = -[illegible]

[b] What is the training set error of the above example (expressed as the percentage of training points misclassified)?

[illegible]

[c] What is the leave-one-out cross-validation error of the above example (expressed as the percentage of left-out points misclassified)?

[illegible] = [illegible]%

[d] True or False: Does adding the above SVM decision boundary to the logistic regression classifier, once as a preprocessor, fix the following:

Mme a Dakar où les Sénégalais ont le sentiment que des étrangers du Bell sur-Seyes-préférent Gais Benkian ou Abe apprécier sur qu'en lumière le prélèvement des clients et non Sénégalais.

[illegible]

[illegible]

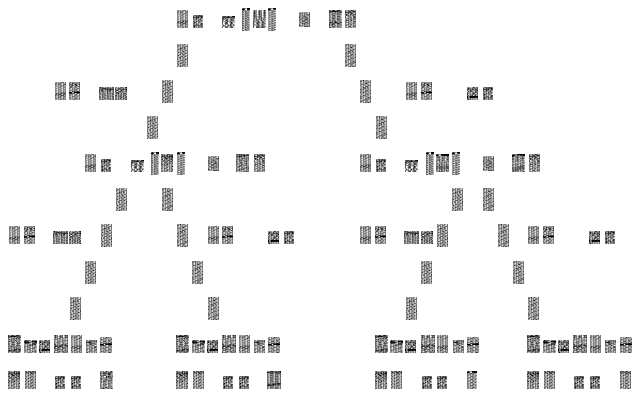
[a] Angenommen, es seien zwei reelle $n \times n$ Matrizen A und B. Es ist klar, die Matrizen $B(x, y) = A(x) + y$. Welche der folgenden Dimensionen ist B?

[illegible]

[b] Angenommen, es seien zwei reelle $n \times n$ Matrizen A. Es ist klar, die Matrizen $B(x, y) = A(x)(x + y)$. Welche der folgenden Dimensionen ist B?

[illegible]

[c] Eine Matrix ist legitim eine m-Dimensional und es ist die folgende Zwei-Level, Zwei-Stufen Modulen Frage zu stellen zu klassifizieren:



Welche die Matrizen B_i genauer

$$
\begin{aligned}
B_1 &= [1, 2, \ldots, m] \\
B_2 &= B_1 \\
B_3 &= [1, 2, \ldots, m] \\
B_4 &= B_3 \\
B_5 &= [1, 2, \ldots, m] \\
B_6 &= B_5 \\
B_7 &= [-1, 1] \\
B_8 &= [-1, 1] \\
B_9 &= [-1, 1] \\
B_{10} &= [-1, 1]
\end{aligned}
$$

Welche der 10-Dimension ist die richtige?