# Del 1

Artificial intelligence:   
Any technique that enables machines to mimic human intelligence.

Machine learning:  
Statistical methods that enable machines to “learn” tasks from data without explicitly programming.

Deep learning:  
Neural networks with many layers that learn representations and task “directly” from data.  
  
From Data to intelligence:  
Data is processed into information  
information is the cognition of knowledge  
Knowledge requires pattern extraction for understanding  
Understanding is inference for intelligence.

# Del 2

Steps of building a machine learning model:  
- Load the data   
- Data preprocessing (Different approaches based on file, missing values and so on)  
- Split the data between training and testing (One to train the data with and one to test the data against)  
- Standardization/normalization (transform into numeric data)  
- choosing and using algorithms  
- Evaluate the model

Missing values  
When it comes to missing values there are multiple approaches based on whats missing:  
- If you simply don’t have the data you could set the data to a default value as 0 or 1.  
- You could also set the value to the median value of all the data combined.  
- You can use the most frequent value that is found.

You also need to remember to clean the data for duplicates (Example if one person have multiple email accounts). You also need to have in mind that scaling are different such as 15kg could mean 0 but 100kg means 1 – Then else were the values are changed. Consistency is key. You could also have inconsistency in the data, as an example number, nr, no and numb and so on. Date formatting is a common failure in datasets.

## Feature selection

The idea of feature selection is to shorten the training time, simplify the model and make it easier to understand. You can also enhance the performance of the model by removing unnecessary data also called overfitting. This is “whitenoice” or data that should not matter/affect/influence the model and can in most cases be removed from the training set.   
For the model to produce accurate and reliable data, you need to make sure it gets the right and correct data. Feature selection will enable your model to make decisions correctly instead of being influenced by data that isn’t necessary.   
The result will make your model faster, more efficient, easier to understand.  
There are different approaches to feature selection which is:  
- IF you understand the data used, you can use machine learning and manually select the data for your model.  
- IF you have limited understanding for the data used, you can try and use automatic selections algorithms such as decision trees, random forests, K-nearest Neighbors(KNN) and so on.  
- IF you have data with a lot of features, you can use component analysis to reduce the dimensions of the data.   
  
There are different techniques for feature selection:  
Feature importance – Give a score of each feature in the dataset, for selecting important features.  
Univariate Selection – Give a score of each feature in the dataset, for determine which feature have the strongest correlation to the output.  
Corrlation heatmap – Provide a matrix to show relationship between the different values of the features.

In feature engineering you create new variables to enhance your model. If the feature does not provide you with the expected results you will have to redo or remove this feature all together.

## Scaling

Scaling is a method to standardize the data of different ranges to a standard method so the data that is different can be combined into a common value for a better dataset.  
You have two different categories:  
Normalization and standardization.  
For normalization you set the values down to 0 and 1, where 0 is lowest and 1 is highest. This is to clean up the data and the understanding around it. As an example there is a lot of difference between data such as flight fuel, icecream consumed and gaming hours. If you normalize the data into the same value from 0 to 1 you could in theory use the data together.  
  
For Standarization you use the same logic only that the data can vary from -1 to +1. So you have a bigger spread of the data where the mean is 0. This can create a different visual and understanding of your data because you can go in both dimensions of the positive and negative side.

# Del 3

## Machine learning

Supervised learning is a task driven learning that has regression and classification as sub-categories. It refers to building the model based on labeled data. Unsupervised is data driven learning with clustering as a sub-category. Unsupervised learning is the process of building an model without relying on labeled data. As an example the supervised data is labeled so the model knows the difference between the data put into the model. You can say that this data is pictures of dogs and the next set is pictures of cats. In unsupervised you simply give the model all your data (pictures) and let it decide for itself the differences. This leads to that the supervised data has a clear separator between the two sets while unsupervised it can only assume based on the cluster of data concentrated in one area.   
  
For supervised data classification it will categorical go through the data and answer questions in simple yes and no (binary method). As an example it will answer yes/no to questions like “Is this a dog?” or “is this a hippo?” and then categorize the data based on these parameter.

For supervised regression it the algorithm predicts the next outcome based on previous data. It will as an example use the general data outcome based on different data parameters to predict its next value. Examples of this can be house pricing, stock market or weather forecasts.  
One could also namedrop reinforced learning, this is simply a reward-like system where the model gets rewarded positively and negatively for doing tasks. Think about a robot traversing a 3x3 grid field and the idea is to let the model step into all the fields in the fastest and most efficient way.

# Supervised

## Algorithms - Classification

In classification it assigns each of the inputs to a given number for discrete categories.

KNN:  
Stands for K-Nearest Neighbor. Which means that it will land on the nearest possible result then say that “since we are so close we are technically the same” kind of deal. So as an example if this result gets a yes, then surely my value will also be yes.   
  
Naïve Bayes:  
This technique is used to build classifiers. It describes the probability of an event happening. This is based on different conditions that are related to the event.  
Usage of this is that its fast and have a relatively high accuracy. The downsides is that when an event is not connected to another event it simply will not work correctly.  
  
Decision tree:  
As the word says it’s a tree that divide the different datasets into different branches. As an example it checks if a target value is higher or lower, if yes it checks another parameter depending on the yes/no in the first branch. The second parameter could be different based on what is checked in the first branch and then it travers down the line like this. Benefits of this is that its very clear at each step it goes through, but if the data is too complex it is very hard for it to create the tree. If its too complex it might use random forest instead.  
  
Random Forests:  
It builds multiple models based on the decision tree method then combine them for a better result. The combination of all these models will create a forest which improves the accuracy and robustness of the model. In theory if the decision tree isn’t enough just run a lot of them to get an entire forest.

## Algorithms - regression

The objective for these algorithms is to find the relation between input variables.   
It aims to find the mathematical relationship between input variables and output values.   
  
Linear regression:  
Predicting a continuous numerical output based on the features input. It models all the relationships between the variables in a linear way.  
  
Logistic regression:  
Used for binary classification problems, mostly yes/no and similar 0/1 outcomes.  
It’s a prediction based on the numeric values.

## Confusion matrix:

A confusion matrix is a tool for understanding the performance of the classification models in deeper detail. It goes beyond the accuracy and provides insight into how well the model is doing with predictions and errors. This information is used in understanding decisions about model selection, parameter tuning, and addressing issues like class imbalance.

# Unsupervised

The idea is to let the machine learn for itself so it may find hidden gems or other patterns. It takes in all the available data without a target. Then it categorizes the data into clusters to make it readable. It discovers inherent categories in the data to the groups.  
Its commonly used to let the model explore data, find outliners and do pattern recognition.  
The model will see things that humans don’t always see.

K-Means:  
This is an algorithm that clusters the data. For it to work the model know the number of cluster beforehand. This is a fixed value it will work up against. It will under each iteration of the model update the center for the clusters, so the clusters move accordingly. The centroid is the name of the center of the cluster that will be placed at their optimal location.

# Del 4

## Differences in supervised vs unsupervised

The distinction between the different approaches is the presence of labeled target data. This value can be set to Y as all featured or attributes in the data will be X. The Y will be found in supervised algorithms while the Y is not a variable in unsupervised algorithms. In supervised you teach the model to make decisions based on labeled data (y) while in unsupervised it explores the data and identify patterns or clusters without a target value.

# Del 5

Ai have a diverse role in various contexts. Beyond just handling household tasks or general daily uses they can be used for chatbots that can be from normal conversations to problem-solving issues. You can also use a typical chatbot to record voice and schedule meetings or other tasks by simply talking to it.

As an example AI can have tasks such as your vacuuming or cut your grass outside, even water your plants in the household. Some AI can even track your refrigerator and what ingredients you are missing. These are called smart home solutions. You can also get them to drive cars or sort wares based on their sizes in packing or post places. They can do calls for you under emergency and they are fitted with commando words it listen to. If you have a google home it will trigger when you say “Hey Google” or whatever the commando word is. This will trigger different effects from the AI and it will respond accordingly.

# Del 6

NLP or Natural language processing. The goal is to get the computer to understand human language. As we humans speak, we communicate something we know, then the listener takes the act/goal of whats being communicated and translate this into a personal understanding and meaning.   
This is something hard for a computer as it only sees the binary system and cannot understand the intent of whats being said. You can teach a machine to understand positive and negative words but it cannot understand the feelings of it.

## Understanding and generating

For the understanding it involves the ability for the AI to understand the human language. For us this means they are taking in words as an example, then translate this word. The issue experienced is that it often translates directly. There are millions of dialects in all languages in the world and the AI does not always understand that. A sentence in Norwegian directly translated into English could output gibberish or hard to understand language. The last part where it tries to output is called generating words and context.

## Processing

The process in this context means how it understand and generate these texts.  
- segmentation: Breaking data into sentences  
- tokenizing: Break sentences into words (this could be step one)  
- stop words: marking down the stop words (commonly used between words)  
- Stemming: Reduce words to its base form  
- Lemmatization: find the base or dictionary form for the words  
- Feature extraction: Transform the text into numerical data  
- Modeling: Build the model with algorithms ex naïve bayes.

## Structured and unstructured data

It can be used for both without issues. The main difference is that in structured data it will often be on a 1-by-1 basis in terms of columns and rows in an excel sheet or similar. While in unstructured data it will read large volumes of text for analyzation and get insight. So the structured data can make the ai understand positive or negative or neutral words. These can often be used as reviews and feedback. While when it reads unstructured data, it cant always predict if the text is positive or negative. Specially when it reads social media texts it can only predict based on a bad word vs a good word. As an example it could say that something very bad is good due to the AI not understanding the context (as in saying WW2 was a good thing).

## How do build a NLP model

Same process as in machine learning:  
load the data, text preprocessing, split the data into training and testing sets, standardization/Normalizing(transform into numeric data 0/1s), Train in naïve bayes, evaluate the model, save the trained model.

# Del 7

## Application of computer vision

Object detection:   
- Identifying and locating objects with images or video frames.  
- Commonly used in cars for recognizing pedestrians, traffic signs and other cars.  
Object tracking:  
- VR headset position tracking, hand and body 3D 360 cameras  
- Monitoring and following movements of objects over time  
- Used for VR, surveillance systems and tracking general movement  
Object verification:  
- Confirming if the object is what’s its expected to be  
- Can be used for facial recognition or finger scanning  
Object identification:  
- Categorizing the detected object into different categories.  
- Can be used for image classification where it can sort pictures into categories such as mushrooms, dog breeds, plants and so on.  
These can be used for picture analysis, face detection, 3D modelling and rebuilding, robotics, optional character recognition(OCR) for pictures with text in them.

You can use convolution neural networks (CNN) which automatically learns and extract hierarchical structures from images.

## Pixels and rgb channels

Feature extraction is one of the crucial steps where the goal is to identify relevant characteristics in an image. These features are then used for further analysis such as object detection, image classification or segmentation.  
These methods can be categorized into handcrafted feature extractions and learned feature extractions using deep learning methods.  
Features in pictures can be color, geometrics such as edges and corners, text extractions and so on. As an example for color it checks the primary color of an animal, it then cross this color code with its internal color code, then search for animals with this color and it can assume based on the layout of the picture what kind of animal this is.  
Pixels are often referred to as attributes for a machine and is used numerical. With all these numbers available for the model it can predict the picture or recreate it in another scale. It can do a very pixilated picture to a useful picture. As an example it can let a picture go from very low quality pixilated to a high quality useful picture. It basically writes every pixel and calculate the need of every single one of them for it to work.   
  
Some algorithms used could be:  
*- SIFT(Scale-invariant feature transform)*  
- SURF(Speeded-Up robust features) Detecting and describing key points in images, faster than SIFT while having high performance.  
*- ORB(Orientated FAST and rotated BRIEF)*

# Del 8

## Neuron tree/Neural networks

Neurons are set up to a tree that starts out with the bottom side with a lot of variables being calculated. They are put into the input layers to be weighted based on probability. It then learns from the weighting and continues to do the backpropagation to update the weighting during the training. After the training is completed it will sit with different possible solutions and will based on the weighting select a solution in the output layer that is the most probable solution. As an example it will go through a lot of neuron nodes, layers, weighting and learning just to figure out if a picture is a t-shirt or a cat.  
  
For simplicity it will “build” up a body based on probability. As an example it will determine if the picture is a horse or a fish based on different parts of the picture. It will then say that the face is 60% probability of a horse and the legs will be 40% probability. If all the probabilities are combined it will create a probability out of that and confirm that it is a horse.  
  
Artificial neural networks “ANN” is a structure that is inspired to function like the human brain.

## Components of neural networks

Neurons: Nodes that process and transmit information  
Layers: Organized into input, hidden and output layers  
Weight and biases: Adjusted during training to minimize errors  
Learning: The process of adjusting weight to improve predictions  
backpropagations: Algorithm for updating weights during training  
Activation functions: Determining if neurons are activated or not.

# Del 9

## Deep learning

CNN(convolutional neural network) are primarily used for image data. The algorithm is designed to do task such as image classification, object detection and image segmentation.  
CNN use convolutional layers to learn features from images which makes them highly used for computer vision tasks. It is commonly used in facial recognition, picture classification, helping doctors with detecting abnormalities in x-rays, diagnose plants and much more.  
  
RNN(recurrent neural network) are designed for preprocessing sequential data. This are texts, speech and voice and more. Its commonly used to voice assistants, suggesting words when typing, speech recognition, predicting patient outcomes and so on. They are basing its decisions on the data coming in. The idea is that the algorithm memorize the previous inputs and use them for the future. This make the algorithm perfect for tasks such as text to speech etc as it previously have encountered the same speech pattern.

They both take in unstructured data as the data can be images, speech or random text on the fly. As an example they are used in the word suggestion on mobile devices that often can be annoying it its always wrong or makes no sense based on your dialect. But they are great tools for those who find them useful.

# Del 10

Current ongoing debates:

Deepfakes, bias and fairness, job replacements, ethical use, customer support, school and education.