

CPE504

Artificial Neural Networks (ANNs)

1.0. INTRODUCTION: MACHINE LEARNING (ML)

What you will learn

- What is Machine Learning (ML)
- Challenges with ML
- Types of ML
- Classification and Regression
- ML Pipeline
- Some Notes on the current State-of-the-Art (SOTA)
- Summary
- Tools
- Recommended Texts

What is ML

What is in a Name?

Generally, the following logical statements are assumed to be universally true

- An **Artificial Neural Net (ANN)** is a kind of **Deep Learning (DL) technology**.
- **Deep Learning** is a kind of **Machine Learning (ML) technology**.
- **Machine Learning** is a kind of **Artificial Intelligence (AI)**.
- Therefore, **Artificial Neural Nets**, the subject of **CPE504**, is a kind of **Artificial Intelligence**.
- ANNs are mechanistic computational abstractions of **Biological Neural Nets** using the universal language of calculus.
- Interestingly, **Artificial Intelligence** is a very common word that may imply many different things. Universally, it is the **representation or imitation** of intelligence **with the aid of computing power**. (**Computational Intelligence**)
- Think: **algorithms (software)** that mimic natural forms of intelligence.

Goal of AI

- To create computer abstractions or models that exhibit **intelligent behaviors** like humans or nature.
- This means machines that can recognize a visual scene, understand a text written in natural language, or perform an action in the physical world.
- **AI** is therefore **intelligence exhibited by computing machines through software**.

What is ML...

- ML itself includes many technologies as well. It is a very generic field of AI that can be found in engineering and the sciences.
- By skillfully **recognizing patterns in data**, deep learning using ANNs have been in the spotlight recently for solving some problems that have challenged AI in an exceptional way.
- It, however, faces the fundamental ML limitations as well. **There are still many things we don't know.**

So, What Is ML?

- ML is a **modeling technique using data**. In control systems engineering and other engineering fields, methods in ML have been traditionally introduced as **System Identification**.
- ML is a term used to represent computational techniques that **identify or learn a model from data**, instead of having a human being generate the model using logical rules or by first-principles.
- **Data** means signals, information such as documents, audio, images, etc.

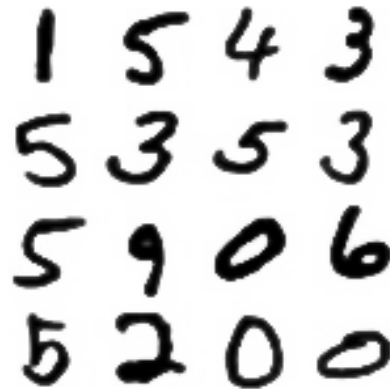
The **final product** of ML or System Identification is a **Model**.

- The data used to learn this model is called **training data**

What is ML...

On Models

- There are many areas where laws and logical reasoning are not very useful for modeling.
- Typical problems can be found where intelligence is involved, such as image recognition, speech recognition, natural language processing, signal processing, etc.
- Machine Learning attempts to solve the problems for which **analytical models are hardly available or are not promising**.
- **Example:** Identify the numbers in the figure below. **Piece of Cake right!**



1	5	4	3
5	3	5	3
5	9	0	6
5	2	0	0

What is ML...

To make a computer do the same thing. What do we do?

- If we use a traditional modeling technique, we will need to find some rule or algorithm to distinguish the written numbers.
- So, **why don't we apply the rules that you have just used to identify the numbers in your brain?** In fact, we see that this becomes a very **challenging problem**.
- Identifying numbers is very easy for humans. Between the **1950s – 1990s**, engineers researching this thought it **must be a piece of cake** for computers to do this, since man-made computers can compute in many ways much faster than humans.
- Well, it did not take very long until **they realized their misjudgment**.

How were you able to identify the numbers without a clear specification or a rule?

- It is hard for you to answer, right?
- **But, why?** It is because we have never learned such a specification.

What is ML...

- From a young age, we have just learned that this is 0, and that this is 1.
- We accepted that information and became better at distinguishing variety of numbers.
- We became very good at **recognizing patterns, imitating others, learning from the accumulation of error and experience.**

What about computers, then? Why don't we let computers do the same thing?

- **That's it! Congratulations!** Machine learning is a way to realize AI.
- You have just comprehended the **concept of Machine Learning (ML).**
- ML was defined in the 1950s by AI pioneer Arthur Samuel as **“the field of study that gives computers the ability to learn a task without explicitly being programmed.”**

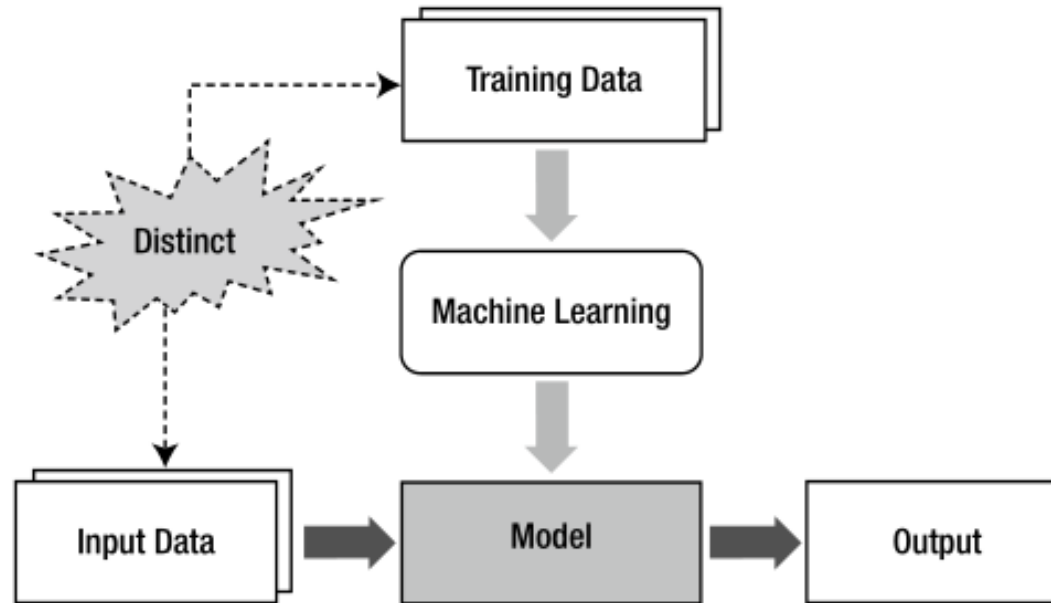
Challenges with ML

- ML is best suited for situations with lots of experience in the form of data.
- So, ML techniques are used to identify (or find or learn) an appropriate model from input data. ML unlike other AI techniques is suitable for problems that involve intelligence, where physical laws, rules, symbolic logic, etc. fail to produce a model.
- ML rules in finding some solutions to such problems. On the other hand, it comes with its own fundamental problems.

Fundamental issues

- First, the ML process finds the model from the training data,
- Then, we apply the model to the test data, simulating the actual real-life data.
- **This process is illustrated in the next figure**, with the **vertical flow** indicating the **learning process**, and the **horizontal flow** describes the trained model indicating **inference**.

Challenges with ML



The distinctness of the training data and input data is the basic root of the challenge faced by the ML process.

- The data that is used in development for modeling in ML and the data supplied in production, that is real-time practical application are most often distinct.

Challenges with ML...

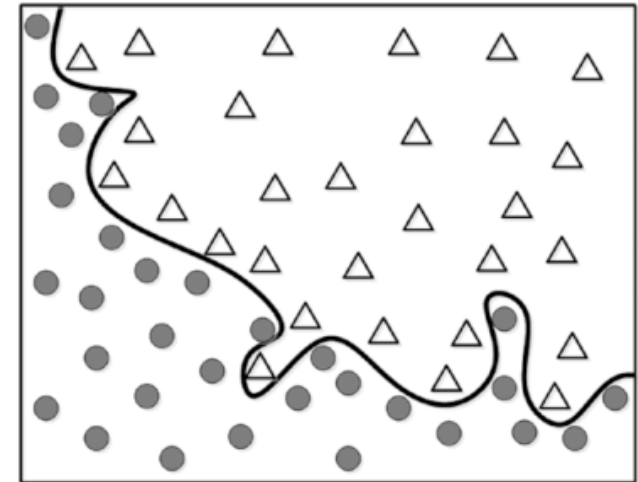
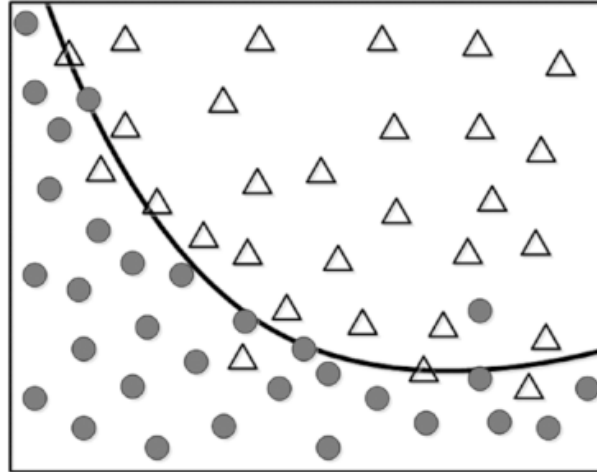
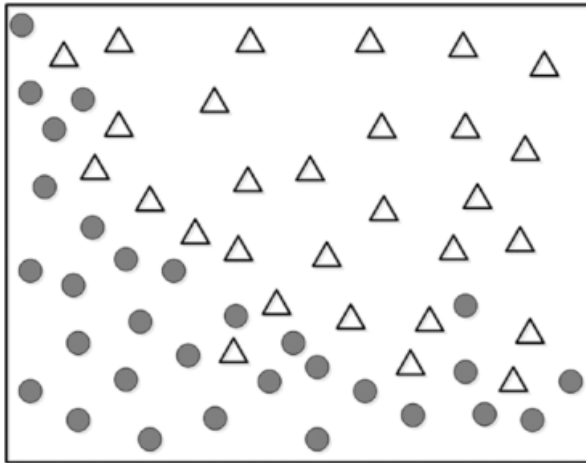
For example:

- **training data:** composed of handwritten notes from a single person.
- **testing data** or **actual field data:** composed of other people's handwritings?
- **Q: End game:** Will the model successfully recognize other people's handwritings?
- **A:** The **possibility** may be very **low**.
- No ML approach can achieve this **desired goal** of **generalization** with the wrong training data.
- Therefore, it is **critical** for ML approaches, especially those to be **deployed in production**, to obtain unbiased training data that adequately reflects the characteristics of the field data.
- **Generalization** is the term used to denote: that the **performance of trained models should be consistent regardless** of the training data during development or input data during production.
- The **success of Machine Learning** relies heavily on how well generalization is accomplished.

Challenges with ML...

Overfitting

- One of the obstacles to improving generalization for trained models is that they may be overfitted to the data. Overfitting is a common problem when we try to fit a ML model to data
- Consider a **classification problem**: shown in the **leftmost** figure below is position data divided into two groups. The points on the figure are the training data.
- The **objective** is to determine a curve (see the **middle** and **rightmost** figures) that defines the border of the two groups using the training data.



Challenges with ML...

- The **middle** figure illustrates what we call a **line of best-fit** described by a model. In contrast, the **rightmost** figure shows an **overfitted line** described by a particular model.
- Real-world **data** or **signal** or **information contains noise**. ML considers all the data, even the noise.
- Since a working ML model of the training data may not reflect the field or production data properly.
- This is why overfitted models, therefore end up producing improper models, and so lower generalizability, when they encounter input data distinct from the training data.
- The knowledge of this does not mean that we should make the model less accurate on the training data on purpose.

Now we face a dilemma!

- reducing the learning error to zero on the training data leads to overfitting.
- overfitting degrades generalizability.

What do we do? The next section introduces certain techniques that can prevent overfitting.

Challenges with ML...

Confronting Overfitting

- Overfitting significantly affects the level of performance of Machine Learning.
- To date, there are two typical methods used to confront overfitting: regularization and validation.

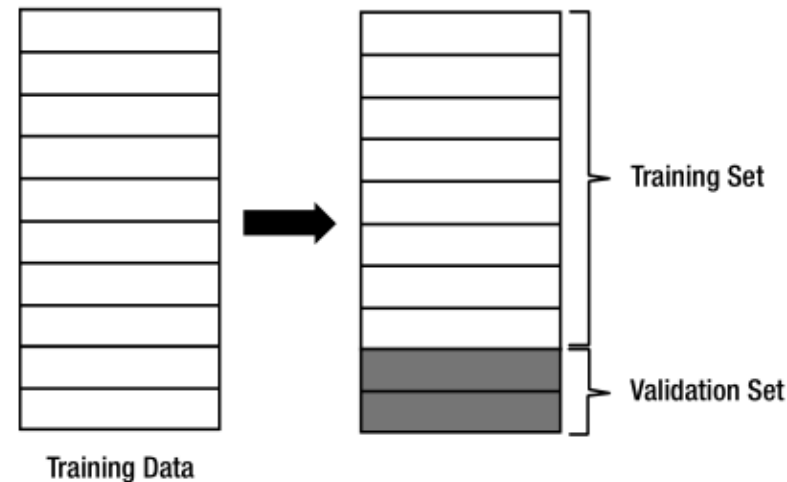
Regularization

- Regularization is a numerical method that attempts to construct a model structure as simple as possible. Such as L1-norm (lasso) and L2-norm (ridge) regularization methods.
- The simplified model can avoid the effects of overfitting at the small cost of performance.
- The complex model tends to be overfitting.
- Using the last example, although it fails to classify correctly some points, the simple curve reflects the overall underlying characteristics common to the group much better.
- Here, we are able to tell that the grouping model is overfitted because the training data is simple, and the model can be easily visualized.
- However, this is not the case for most situations, as the data has higher dimensions. We cannot draw the model and intuitively evaluate the effects of overfitting for such data.
- Therefore, we need another method to determine whether the trained model is overfitted or not.
- This is where **validation** comes into play.

Challenges with ML...

Validation

- Validation is a process that reserves a part of the training data and uses it to monitor the trained model performance.
- The validation set is not used for the training process.
- Because the modeling error of the training data fails to indicate overfitting, we can use some of the training data to check if the model is overfitted.
- We can say that the model is overfitted when the trained model yields a low level of performance to the reserved data input. In this case, we will modify the model to prevent the overfitting.
- The next figure, illustrates the division of the training data for the validation process.



Challenges with ML...

When validation is involved, then:

the training process of Machine Learning proceeds by the following steps:

1. **Divide** the training data into two groups or sets:

- one set for training and the other set for validation.
- as a rule of thumb, the **ratio** of the training set to the validation set is **0.8 : 0.2**.

2. **Train** the model with the training set.

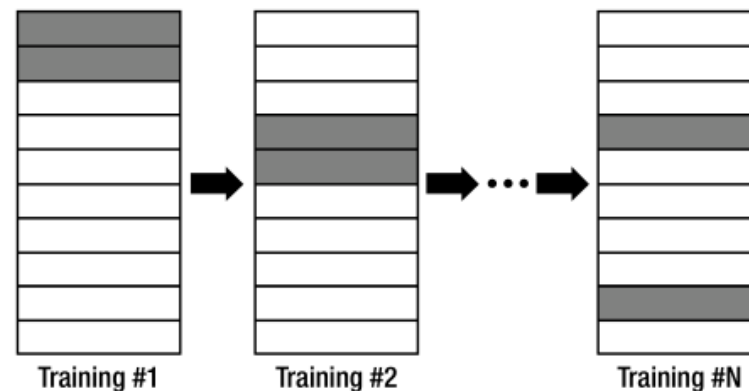
3. **Evaluate** the performance of the model using the validation set.

- a. If the model yields satisfactory performance, **finish the training**.
- b. If the performance does not produce sufficient results, modify the model and **repeat the process from Step 2**.

As there is no golden rule for this **ratio**, a often common and recommended type of the validation process, especially in statistical science is **cross-validation**. There are different types such as k-fold cross-validation. They are all re-sampling techniques.

Challenges with ML...

- **Cross-validation**, like the vanilla validation process, divides the training data into sets for the training and validation, but it keeps changing the sets, instead of retaining the initially grouped sets.
- In other words, cross-validation repeats division of the data.
- The reason for doing this is that the model can be overfitted even to the validation set when it is fixed. As **cross-validation maintains the randomness of the validation dataset**, it can better help us **detect the overfitting** of the model being trained.
- The **figure below**, better describes the concept of cross-validation. The dark shades indicate the validation set, which is randomly selected throughout the training process.



Types of ML

The **function** of a **machine learning model** can be

- **descriptive**, meaning that the model uses the data to explain what happened;
- **predictive**, meaning the model uses the data to predict what will happen; or
- **prescriptive**, meaning the model will use the data to make suggestions about what action to take.

ML techniques can be classified into three general types **depending on the training**.

1. **Supervised learning**
2. **Unsupervised learning**
3. **Reinforcement learning**

Types of ML...

Supervised learning

Supervised learning, whether **implicit or explicit**, is **fundamentally**, the **process by which humans learn**.

Consider that humans obtain new knowledge as we solve exercise problems.

1. **Select** a problem.
2. **Apply** current knowledge to solve the problem.
3. **Compare** the answer with the known solution.
4. **If the answer is wrong**, modify current knowledge.
5. **Repeat** Steps 1 to 3 for all the problems.

ML process: exercise problems and solutions correspond to the **training data** { input, correct output }

The **knowledge** corresponds to the **model**.

The important distinction here, is the fact that we need the solutions. This is the vital aspect of the supervised learning.

In supervised learning, each training dataset should consist of an input and a corresponding correct output pairs.

Learning in this case, is then the series of corrections to a model so that, for a given input, the difference between the correct output and the output from the model is reduced. If a model is perfectly trained, it will produce a correct output that corresponds to the input from the training data.

Types of ML...

Unsupervised Learning

- In contrast, the training data of the unsupervised learning contains only inputs without correct outputs
- Training data structure: { **input** }
- **Unsupervised learning** is generally used for investigating or pre-processing the characteristics of the data or signal.
- This concept is similar to a someone who just sorts out problems by features and doesn't learn how to solve them because there are no known correct outputs.

Reinforcement Learning

- **Reinforcement learning** employs sets of input, some output, and a grade (reward) as training data.
- It is generally used when optimal interaction is required, such as control and game theory.
- Training data structure: { **input, some output, grade for this output** }

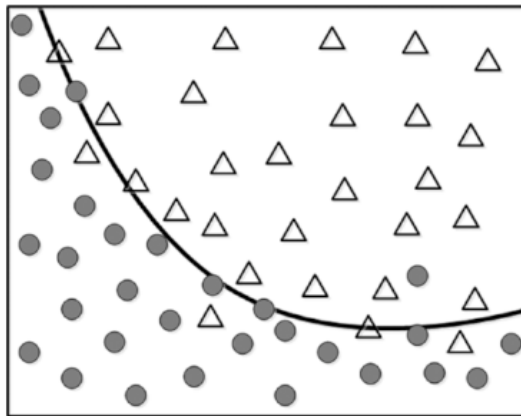
CPE504 only covers supervised learning.

- This form is used for more applications compared to the other types, and more importantly, it is a **fundamental concept in ML**.

Classification and Regression

The two most common types of application of supervised learning are classification and regression.

- The classification problem focuses on literally finding the classes to which the data belongs. Some examples may help.
- Spam mail filtering service: Classifies the mails by regular or spam
- Digit recognition service: Classifies the digit image into one of 0-9
- Face recognition service: Classifies the face image into one of the registered users

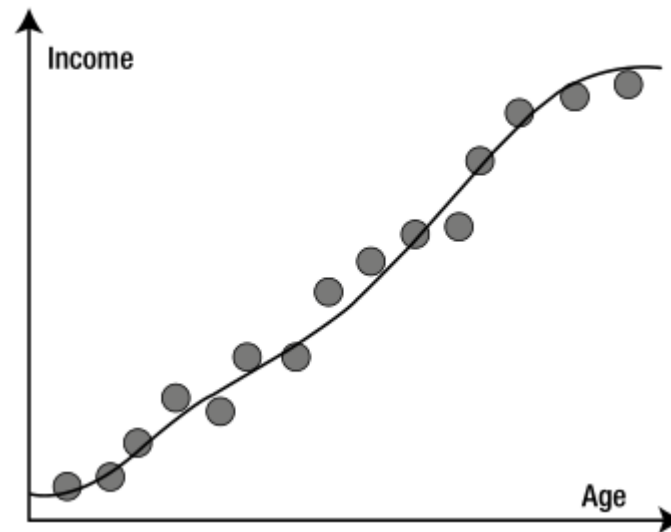


Classification and Regression

- We addressed in the previous section that supervised learning requires input and correct output pairs for the training data. Similarly, the training data of the classification problem looks like this: { input, class }
- In the classification problem, we want to know which class the input belongs to. So the data pair has the class in place of the correct output corresponding to the input
- In contrast, the regression does not determine the class. Instead, it estimates a value. As an example, if you have datasets of age and income and want to find the model that estimates income by age, it becomes a regression problem (see Figure)
- The dataset of this example will look like the table in Figure 1-16, where X and Y are age and income, respectively.

Classification and Regression...

- Regression here, means going back to a set of average values.
- Both classification and regression are parts of supervised learning.
- Therefore, their training data is equally in the form of {input, correct output}.
- The only difference is the type of correct outputs—classification employs classes, while the regression requires values.
- In summary, an ML application becomes classification when we need a model to judge which group the input data belongs to and regression when the model to estimate the trend of the data.



Classification and Regression...

- Just for reference, one of the representative forms of **unsupervised learning** is **clustering**.
- Clustering investigates the characteristics of the individual data and categorizes the related data.
- It is very easy to confuse clustering and classification, as their results are similar.
- Although they yield similar outputs, they are two completely different approaches.
- We have to keep in mind that clustering and classification are distinct terms.
- When you encounter the term **clustering**, just remind yourself that it focuses on **unsupervised learning**.

ML Pipeline

1. **Define** a problem
2. **Prepare** a large set of unbiased, clean data
3. **Train** and **Evaluate** ML model(s) on the data
4. **Finalize** the ML model development
5. **Deploy** the ML model in production
6. **If problems, Go back** to step 1, 2, or 3.

Some Notes: SOTA in ML

- Today, Machine learning is core technology in the business models of many big-tech companies, including Google, Microsoft, Facebook, Amazon, Netflix.
- Usage span: Recommendation engines, Search and Translation, Image analysis and object detection, Automatic chat-bots, Self-driving, Medical imaging and diagnostics, etc.
- While machine learning is over-hyped, there are several things we should know about its current limits.

Explainability

- One area of concern is the ability to be clear about what the ML models do and how they make decisions. This is especially important because systems can be fooled and undermined, or just fail on certain tasks, even those humans can perform easily.
- For example, adjusting the metadata in images can confuse computers or with a few adjustments, a machine identifies a picture of a dog as a cat.

Some Notes: SOTA in ML

Bias in Data

- Machines are trained by humans, and human biases can be incorporated into ML models through data that reflects existing inequities, and therefore help perpetuate forms of discrimination.
- For example: Chatbots trained on how people converse on Twitter can pick up on offensive and racist language, for example.
- Also Facebook Ads used machine learning as a tool to show users ads and content that will interest and engage them — which has led to models showing people extreme content, spread of conspiracy theories or inaccurate content.
- This is why data (signal or information) processing or handling is fundamental in ML.

Computational Overhead

- The heavy computational demand of ML algorithms especially for DL leads to the need for large scale compute power (GPUs and TPUs). Currently the possession of these machines can mostly be found in big-tech companies or big-name research groups, which they have used to demonstrate impressive results.

Some Notes: SOTA in ML

Tool Democratization

- There has been a somewhat worrisome push, since 2015 by mainstream big-tech companies and certain class of programming communities, for others, to adopt and use their own developed libraries, automation tools, frameworks, and so on to do ML. Poster examples are: companies like: Google, Facebook, Microsoft, IBM, Nvidia, Intel, MathWorks, etc., and also general-purpose programming languages like Python.
- This has made it a case of knowing which tools to use, in business settings.
- *The truth however is that any other tool or languages can be used, and since the 1950s, there exist many implementations of ML algorithms in different languages and tools.*

Evolving Knowledge and Artistry

- Take a look at:
- <https://twitter.com/tdietterich/status/1374104340613324802>
- <https://arxiv.org/abs/2105.01601>

Some Notes: SOTA in ML

- <https://arxiv.org/abs/2105.07576>
 - <https://arxiv.org/abs/2105.02723>
 - <https://twitter.com/nicvadivelu/status/1390754518481215492>
 - https://twitter.com/Hanxiao_6/status/1394742841033641985
 - <https://arxiv.org/abs/2105.08050>
 - https://twitter.com/MIT_CSAIL/status/1388885995328843776
 - <https://twitter.com/fchollet/status/1386373123889528835>
 - <https://arxiv.org/abs/2102.06171>
 - <https://twitter.com/elonmusk/status/1387901003664699392>
 - <https://arxiv.org/abs/2008.07970>
- *All these are evidence that the tallest trees in ML haven't been fell yet. There are still many things we don't know. Also there are many things we think we know but really, we don't know.*

Summary

Terms we have learnt:

- **Machine Learning (ML)**

Machine Learning induces a model from training data. **ANN is DL. ANN is ML. ANN is AI.**

Data fitting, Compression, Statistics, Optimization,

- **Overfitting**
- **Regularization**
- **Validation**

- **Generalization**

Due to the differences between the training data and actual input data, we need a sufficient amount of unbiased training data, to ensure better generalization.

- **Types of ML based on training data format or structure**

Supervised learning can be divided into classification and regression, depending on the usage of the model.

- Classification determines which group the input data belongs to. The correct output of the classification is given as categories.
- In contrast, regression predicts values and takes the values for the correct output in the training data.

Tools

Recommended Languages

- MATLAB
- JavaScript

Instructions to Student

- **All ANN functions will be written from scratch in form of custom libraries**
- **Learn to transfer maths to software.**
- **Copying of another person's code (or work) will be heavily penalized.**

Recommended Texts

Main Texts

- **MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence** by Phil Kim
- **Pattern Recognition and Machine Learning** by Christopher M. Bishop
- **Understanding Machine Learning: From Theory to Algorithms** by Shai Shalev-Shwartz and Shai Ben-David
- **PATTERNS, PREDICTIONS, AND ACTIONS: A story about machine learning** by Moritz Hardt and Benjamin Recht
- **Mathematics for Machine Learning** by Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong

Good luck!