# Artificial Intelligence & Machine Learning
## 'Bridging the Skills Gap'

**Intro**

Hello, my name is Mario Favaits and welcome to my YouTube series on Artificial Intelligence Machine Learning (ML).

In this session we will answer following questions:

1. Why this video series and why now? What is the skills gap? How do we quantify the skills gap? What is the opportunity for viewers?
2. How is this series different from other YouTube series on the same subject?
3. What is the target audience of this series and what are the prerequisites for this course in terms of math and coding skills?
4. How much effort will it take to become AI and ML savvy?
5. What AI/ML subjects will the series cover?
6. How many videos will be uploaded per month?
7. What are the reference materials used for this series?
8. What software and tools will I need in order to be able to get maximum value from this video series?

*Why this video series and why now?*

If we look around us, we can detect a <u>hype</u> with respect to everything that has to do with digitalization. Don't get me wrong, the digital revolution has started, and it will have a big impact on the way we do things but…people tend to be overly optimistic/pessimistic mainly because a lack of knowledge about the subject.

On a daily basis we read/hear about:
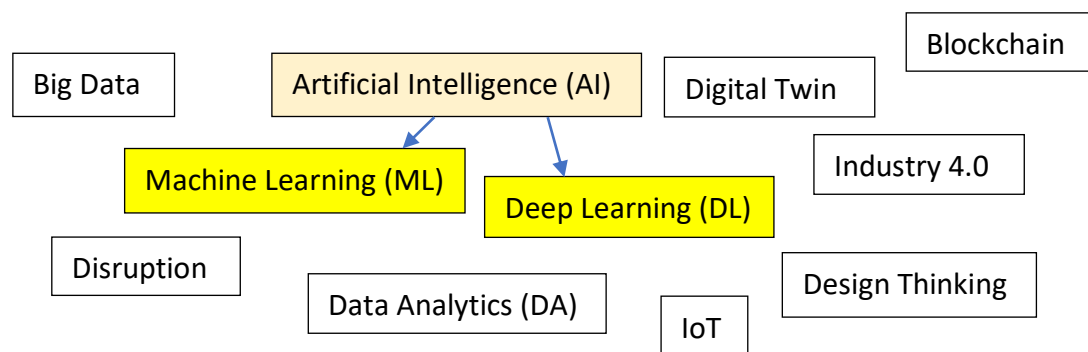


In this video series we will only deal with subsets of AI: ML and DL

*Side effects of the digital hype:*
1. Everybody talks about AI and ML but…most people don't know how to do it!
2. How big is 'BIG DATA'? Lots of companies do not have '*BIG DATA*' $\rightarrow$ it becomes important to understand how to apply AI and ML on '*smaller*' data sets!
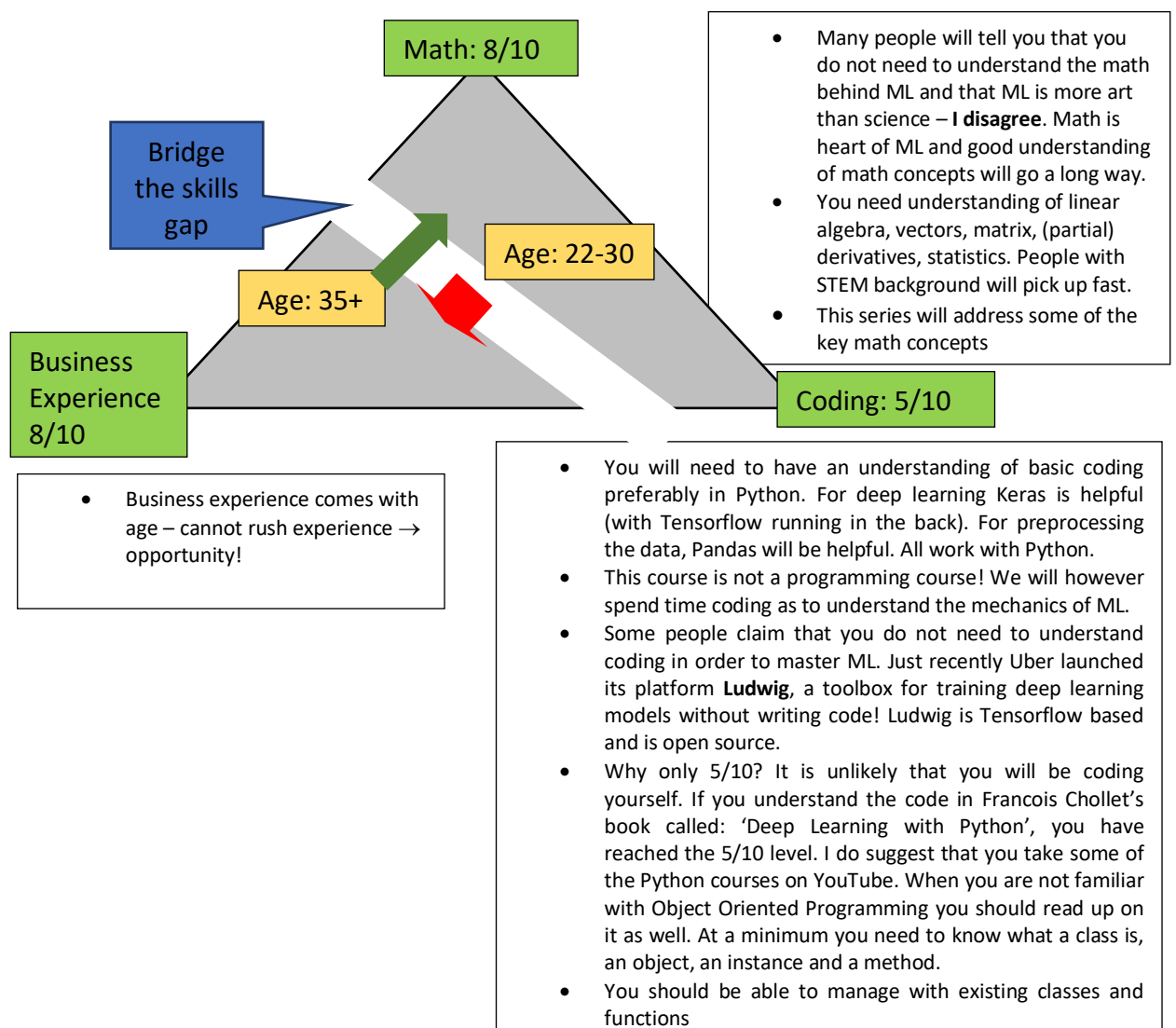
3. Before companies are able to effectively use AI and ML, they will need to go through a serious data governance exercise. Data needs to be correct, complete and in the right format before it can be worked on by AI and ML models. Companies underestimate the backend plumbing that is required to have an effective AI/ML application
4. Companies tend to believe that AI and ML will solve all their problems!
5. Universities (Harvard, MIT, NUS) came out with a curriculum addressing many of the subjects mentioned above:
    a. traditional curriculum – you are basically looking at 2-4 years of study
    b. executive courses – short, incomplete and expensive
6. Software vendors including Google, Amazon, SAP, Microsoft offer cloud-based ML modules as part of their business software → software waiting to be used!

This series aims at clearing the mist with respect to AI and ML and will allow you to build a solid and hands-on foundation on the subject.
It will give business leaders the skills and knowledge to maximize the value of their data, manage data projects more effectively and inspire and hire the best data scientists and

*What is the skills gap? How do we quantify the skills gap? How will this series create an opportunity for viewers?*

In my opinion you need **three attributes** (Mario's triangle)

Math: 8/10

Bridge the skills gap

Age: 22-30

Age: 35+

Business Experience 8/10

Coding: 5/10

- Many people will tell you that you do not need to understand the math behind ML and that ML is more art than science – **I disagree**. Math is heart of ML and good understanding of math concepts will go a long way.
- You need understanding of linear algebra, vectors, matrix, (partial) derivatives, statistics. People with STEM background will pick up fast.
- This series will address some of the key math concepts

- Business experience comes with age – cannot rush experience → opportunity!

- You will need to have an understanding of basic coding preferably in Python. For deep learning Keras is helpful (with Tensorflow running in the back). For preprocessing the data, Pandas will be helpful. All work with Python.
- This course is not a programming course! We will however spend time coding as to understand the mechanics of ML.
- Some people claim that you do not need to understand coding in order to master ML. Just recently Uber launched its platform **Ludwig**, a toolbox for training deep learning models without writing code! Ludwig is Tensorflow based and is open source.
- Why only 5/10? It is unlikely that you will be coding yourself. If you understand the code in Francois Chollet's book called: 'Deep Learning with Python', you have reached the 5/10 level. I do suggest that you take some of the Python courses on YouTube. When you are not familiar with Object Oriented Programming you should read up on it as well. At a minimum you need to know what a class is, an object, an instance and a method.
- You should be able to manage with existing classes and functions

*How is this series different from other YouTube series on the same subject?*
The series will be comprehensive (one stop shop including math and coding) and will review the most relevant topics of AI and ML in detail allowing people not familiar with the subject to become 'experts' **(8,8,5)** in 6 months.

*What is the target audience of this series and what are the prerequisites for this course in terms of math and coding skills?*
This YouTube series focusses on bridging the skills gap allowing business leaders to become skilled with respect to AI and ML. This series will go into quite an amount of detail and it will enable you to add value relatively fast. Having a **STEM** (Science, Technology, Engineering and Math) background will definitely help. However, I am convinced that you can get through these series even when both your math and coding foundation is weak. It will just take that little bit more effort and dedication.

*How much effort will it take for me to become AI and ML savvy?*
My estimate is that you will have a solid understanding of AI and ML if you allocate an average of 15hrs a week for roughly 6 months depending on your foundation in math and coding.

*What subjects will this series cover?*
- Math refreshers: Linear Algebra, Calculus, Statistics
- AI and ML History and Basic Concepts
  - Supervised vs. Unsupervised learning
  - Regression - Classification
  - Loss/Cost Function
  - Gradient Descent/Stochastic Gradient Descent/Mini-Batch Stochastic Gradient Descent
  - Bias and Variance (under- and overfitting)
  - Regularization
  - Feature Scaling
  - Optimizers (Momentum, Newton, Adagrad, RMSprop, Adam)
  - Dimensionality Reduction – Principal Component Analysis (PCA) and Singular Value Decomposition (SVD)
  - Vanishing and Exploding Gradient
  - Bagging & Boosting (Adaboost)
- Classification Algorithms
  - K-Nearest Neighbors
  - Perceptron - Activation Function
  - Logistics Regression
  - 'Naïve' Bayes
  - Binary Classification
  - Multi-Class Classification - Softmax
  - Support Vector Machines and Kernel Functions
  - Decision Trees
  - Neural Networks

- Neural Networks:
  - Convolutional Neural Networks (Covnets)
  - Recurrent Neural Networks
    - Continuous Bag of Words (CBOW)
    - Skip-Gram
  - Backpropagation
  - Dropout
  - Pooling
  - Batch Normalization
  - Word embeddings: Glove and Word2Vec
  - STML
  - Transfer Learning
- Unsupervised Learning
  - Clustering – K-means
  - Anomaly Detection
  - Recommender Systems
- …

*How many videos will be uploaded per month?*
Plan is to upload a minimum 4 videos per month. Videos are of different length and a few videos will be dedicated to math concepts/proofs. The series will be modular enough in order to fit everybody's math/coding levels.

*What are the reference materials used for this series?*
There is plenty of material available on the web and as usual there are some good ones and some bad ones.

The basis for the sessions will be:

***'Deep Learning with Python'* (Author: Francois Chollet).**
It provides a very detailed and comprehensive overview of Deep Learning (DL), keeping math content to an absolute minimum. However, the coding in Francois' book is challenging when you don't have any coding experience. Doing the coding exercises however will be very helpful for a deeper understanding of how DL algorithms work. Understanding the code in Francois' book will bring you to a level **5/10**. The ability to tweak the code in Francois' book will bring you **> 5/10**. For example, using pandas to preprocess data instead of python's shutil and os libraries

The YouTube series on ML and DL from **Andrew Ng** is also very useful as it provides a good overview of the majority of topics. The math however is challenging and sometimes confusing.

**Kilian Weinberger** (Cornell) offers a YouTube course on supervised (shallow). This course is mathematically challenging especially with respect to statistics. DL is not discussed in detail.

There are also specific papers of some models that can be accessed online

*<u>What software and tools will I need in order to be able to get maximum value from this series?</u>*

We will be using **python**, **keras** and **pandas** for data preprocessing and coding. However you do not need to install any software because we will be using Google's cloud-based Jupyter-style **Colab** interface: https://colab.research.google.com/notebooks/welcome.ipynb
Through Colab you will have access to **GPU** and **TPU**. You will need to setup a Google account to access Colab!

It is recommended that you open an account at **GitHub** and **Kaggle**. Kaggle was created in 2010 and organizes ML competitions and has lots of datasets that you can play with. All documents, notes, and code projects will be made available to you via GitHub: https://github.com/mfavaits/YouTube-Series-on-Machine-Learning

Most of the course notes are not typed out and when it is the case, I will scan the notes and post them on GitHub.

Interaction with you will mainly be by whiteboard sessions and for the coding we will use screen recording. Videos will be kept as short as possible and we will avoid waiting time for whiteboard cleaning etc.

**Next session:**

In the next session we will address following items:

1. What is the difference between Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL)?
2. History of AI/ML/DL
3. Definitions & Basic concepts of ML/DL (Loss/Cost Function, Mean Squared Error (MSE), Regression, Classification, Supervised Learning, Unsupervised Learning)

I will also upload math refresher videos for Linear Algebra, Calculus and Statistics.

*Please do not forget to subscribe to this series and give it a thumbs up when it adds value to you. Do not hesitate to send me feedback/comments on the course so that I am able to improve the quality of the series as we go along.*