I took some notes from **Cihan Özhan's video** **https://youtu.be/ynh3OR_bmns**

It was my homework from **Safebox Cyber Security Data Science Course**

Basics of AI

AI is a production of biomimetic. Biomimetics or biomimicry is the emulation of the models, systems, and elements of nature for the purpose of solving complex human problems.

For example, airplanes were inspired by birds who fly up high in the skies.

You need aerodynamic scientists to work on airplanes, however it is not enough to mimic the entire nature just with one field's scientists. There are countless areas, thus, biomimetics is a multidisciplinary work.

Understanding how AI works, requires understanding the human brain, and it is exactly how it happened at first when AI was just a thought in 1956. In 1959, Ord. Prof. Dr. Cahit ARF gave a seminary about "Can Machines Think and How?" started the AI in Türkiye first. As the question is open ended, there are many definitions for artificial intelligence. There is not a single definition, yet there are good ones. For example,

"Artificial intelligence is a set of theories and techniques that develop complex computer programs that are able to simulate certain traits of human intelligence (reasoning, learning, etc.)." [1]

With the inventions in neuroscience, it was a really strong thought to mimic the neural networks to create a new brain.

Is AI Dangerous?

There is a conflicting thought where "to help humanity, you should destroy humanity" logic, and if one day a Super AI finds it true, we might extinct. However, in today's media, the news that you see such as "Google AI created its own *baby AI*" does not really represent the truth.

Sources to Study AI

Andrew Ng is one of the heroes of AI, Andrew Ng is Founder of DeepLearning.AI, General Partner at AI Fund, Chairman and Co-Founder of Coursera, and an Adjunct Professor at Stanford University. Here are some resources: [2]

1- https://www.coursera.org/collections/machine-learning
2- https://www.coursera.org/instructor/andrewng

Stages of AI

1) Narrow AI (ANI): Dedicated to assist with or take over specific tasks.
2) General AI (AGI): Takes knowledge from one domain, transfers to another domain.
3) Super AI (ASI): Machines that are an order of magnitude smarter than humans.

And to be real, it is only the ANI that exists in our world yet. While it is questionable to create an AGI, it may not be so far from today. Today’s chat-bots, social media interaction AIs, personnalized studies or ads, chess bots etc. are all ANI and can do only their job and nothing more. AGI and upper level are like human, capable of learning different subjects and connecting them together to make sense. But in our situation, without even completely knowing the human’s decision system, it is hard to mimic a better decision system.

AI Winter

There were many interruptions in the development of AI due to various reasons, mostly about financing researches and limited physical equipment. Back then, today’s strong CPUs and GPUs did not exist.

In 1960s there were inventions, yet Convolutional Neural Networks (CNN) was not applicable back then.

In 1980s, a new algorithm called back-propagation was found and CNN was popular once again.

In the end of the 1990s, Support Vector Machines and Kernel Trick came, and CNN was gone again.

We are currently on one of the longest periods of AI researches in AI history.

In 2009, Geoffrey Hinton and his students developed a new training method to speech recognition.

Geoffret Hinton and his student Alex Krizhevsky, developed a 7-layered CNN, using the model in 2009. [3]

Dropout was developed to stop overfitting. Overfitting means memorising everything when it comes to the point, if AI sees anything a bit different than the training data, it will not detect as it should.

In 2012, Computer Vision era started with ILSVRC 2012 contest where there were 1.2 million images and 1000 classes. The main problem was to detect an image and come up with 5 estimations. If one of them is true, the model is found to be successful.

The winter never came back again.

Opportunities and Threats

AI could replace many of the industry, including Law System, Drivers, Medical Doctors… anything that requires studying of text or anything automated.

However, although AI is expected to be heuristic, it is not much developed yet. The AI cannot learn everything on its own, there is still need for guidance.

Scientific Basis of AI

Computer Science in hardware and software,
Philosophy for decision,
Mathematic for logic, algorithm, and optimization,
Psychology for modeling human thinking process,
Neuroscience and Biology for understanding the process in low level and mimic it,
Linguistics for designing programs
are needed, once more mentioning this is a multidisciplinary process.

Subtopics of AI

Artificial Intelligence: The development of computer systems capable of performing tasks that typically require human intelligence, such as perception, reasoning, learning, and decision-making.

Machine Learning: A subset of AI focused on developing algorithms that enable computers to learn from data and improve their performance without explicit programming.

Deep Learning: A specialized branch of Machine Learning that constructs artificial neural networks with multiple layers to process complex patterns and representations from large datasets.

Data Science: An interdisciplinary field combining statistics, mathematics, computer science, and domain knowledge to extract insights and knowledge from structured and unstructured data.

Computer Vision: A field within AI that enables computers to interpret and understand visual information from images or videos, enabling tasks such as object detection and image classification.

Natural Language Processing: A subfield of AI that equips computers with the ability to understand, analyze, and generate human language, facilitating tasks like sentiment analysis, language translation, and speech recognition.

Data Engineering: Involves designing and managing systems and processes to collect, store, and process large volumes of data efficiently and effectively, often in conjunction with Machine Learning applications.

Gaming

Video games are a huge field, as there are billions of people interested in the world. It keeps growing and the application of AI is unavoidable.

AI has several intentions on gaming:

Making smarter gamers,

Avoiding *Churn Prevention* which results in stop of gaming due to boredom of the player.[4]

Better Matchmaking

Automated QA and GamePlay Test Bot that allows testing it faster, replacing test engineers.

Fraud/cheat Detection

…

Cyber Security

• Spam Filter Applications (spamassassin)

• Network Intrusion Detection and Prevention

• Fraud Detection

• Credit Scoring and Next-Best Offers

• Anomaly Detection

• Botnet Detection

• Secure User Authentication

• Cyber Security Ratings

• Hacking Incident Forecasting

Malware Analysis

Analysing how a program acts in a sandbox, or without executing the program, analysing the code with AI.

Some Algorithms

Random Forest [5]
Decision Tree
Support Vector Machines
Deep Learning

Swarm AI

It is a matter of creating a collective brain. It was inspired by the nature of bees.

Preparation for an AI project

Defining the Problem

Finding the domain expert to work with

Defining the Goal

Finding datasets

Using the correct tech

The DevOps Process

Pipeline:

1- Prepare Data
2- Train Model
3- Package
4- Validate
5- Deploy
6- Monitor

The key point is the data must be clean. In real life, us, humans tolerate noises and focus on the real matter, however, in AI there are no noises. Every pixel counts and the data should be fine labelled and balanced. For example, in a vehicle detector AI, there should be equal numbers of trucks and race cars. If one of them is too many, the AI could learn as “the others are not much of a car”, resulting in a different way of overfitting failure.

AI can be coded in any programming language yet there are specific ones, that facilitate us in different ways. Python 3 is the best language nowadays; it is easy to find the community to help and it keeps getting updated. Back in the day Java was in use, and in some companies still so. It may still be good if your team has a vast knowledge in Java but to start a project, python is a better option

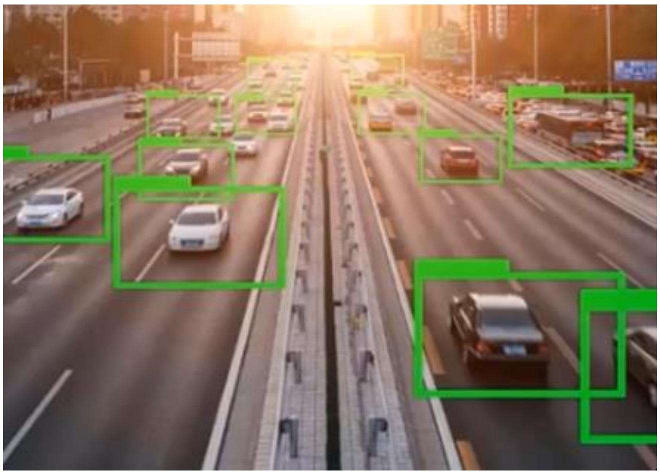
You can also use R if your background is finance, or you excel at Excel.

NVIDIA CUDA GPUs

Nvidia used to create strong GPUs mostly targeting gaming and modelling. However, now the AI and Blockchain have lots of investments, there are specific GPUs for those areas.

…

Labelling



There are opensource labelling tech you can facilitate from:

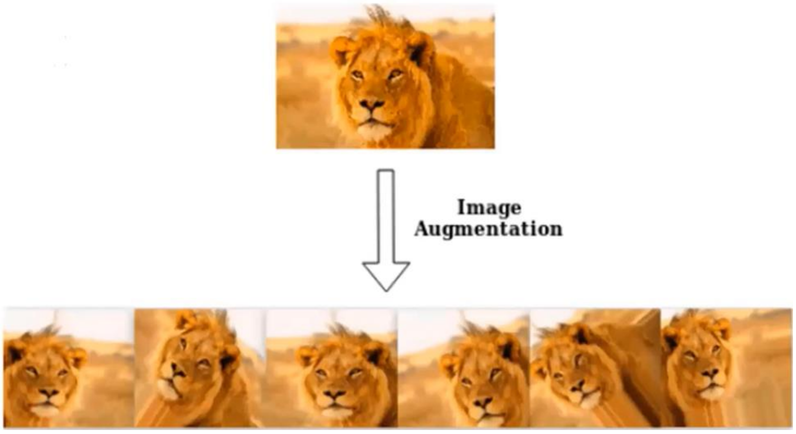
https://github.com/heartexlabs/awesome-data-labeling

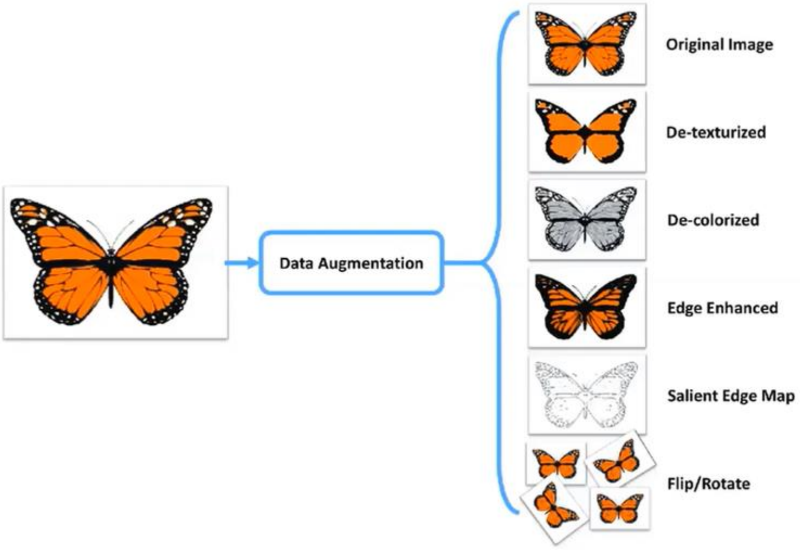
https://github.com/jsbroks/awesome-dataset-tools

https://github.com/rodrigofay/awesome-data-labeling

Data Augmentation

If your data is not enough, or messy that causes overfitting, you can augment your data with some tools. For image augmentation:





There are many DataSets to find on the Internet, yet the best might be Kaggle.

https://www.kaggle.com/datasets

Free AI Starter Kit

Google CoLab allows you to use their GPU to train your program.

There is also Microsoft Azure Notebook, but it comes with CPU support.

References

[1] Marvin Minsky Definition of AI https://link.springer.com/chapter/10.1007/978-3-030-21445-6_2

[2] Andrew Ng https://www.coursera.org/instructor/andrewng

[3] AlexNet https://learnopencv.com/understanding-alexnet/

[4] Churn Prevention https://towardsdatascience.com/how-to-leverage-ai-to-predict-and-prevent-customer-churn-f84d653a76fb

[5] Random Forest https://www.ibm.com/topics/random-forest