ECOM6023

eFinancial Services

LECTURE 8:
Artificial Intelligence and Machine Learning

Dr. Juergen Rahmel 2020-21

HSBC Germany / Hong Kong
IETC Information Engineering – Teaching and Coaching Ltd.

L8 – AI and ML

Topics

- **The Problem Statement**
- Components of AI
- Application Scenarios in Banking and Risk
- Ethical Governance of AI -> separate Slide deck

L8 – AI and ML

The Problem Statement

- What is Artificial Intelligence?
- How can we tell?
- How can we create?
- What are the building blocks?

Artificial Intelligence is made up of three basic concepts:

(i) Capabilities
(ii) Knowledge
(iii) Learning

Cognitive systems will differ in their realization of (i), (ii) and (iii), thus being better at some tasks than others

L8 – AI and ML

The Problem Statement

- Remember the Movie ‘War games’? (it’s from 1983…)
- At last, the computer played tic-tac-toe against itself – and never won. It drew a conclusion from that…
- Tic-tac-toe: 26,830 different games. Full search is easy.
  - There is a ‘perfect’ strategy that secures at least a tie
  - if you can’t win, what’s the point of playing?
- Needed for game playing:
  - A representation of the board and allowed moves
    - i.e. possible board states (configurations) and state transitions (moves)
  - A strategy to win
    - i.e. an evaluation of states and its optimization through selected moves
- What about other games? Chess? Go?
- How do Humans play games? By enumerating possible states?

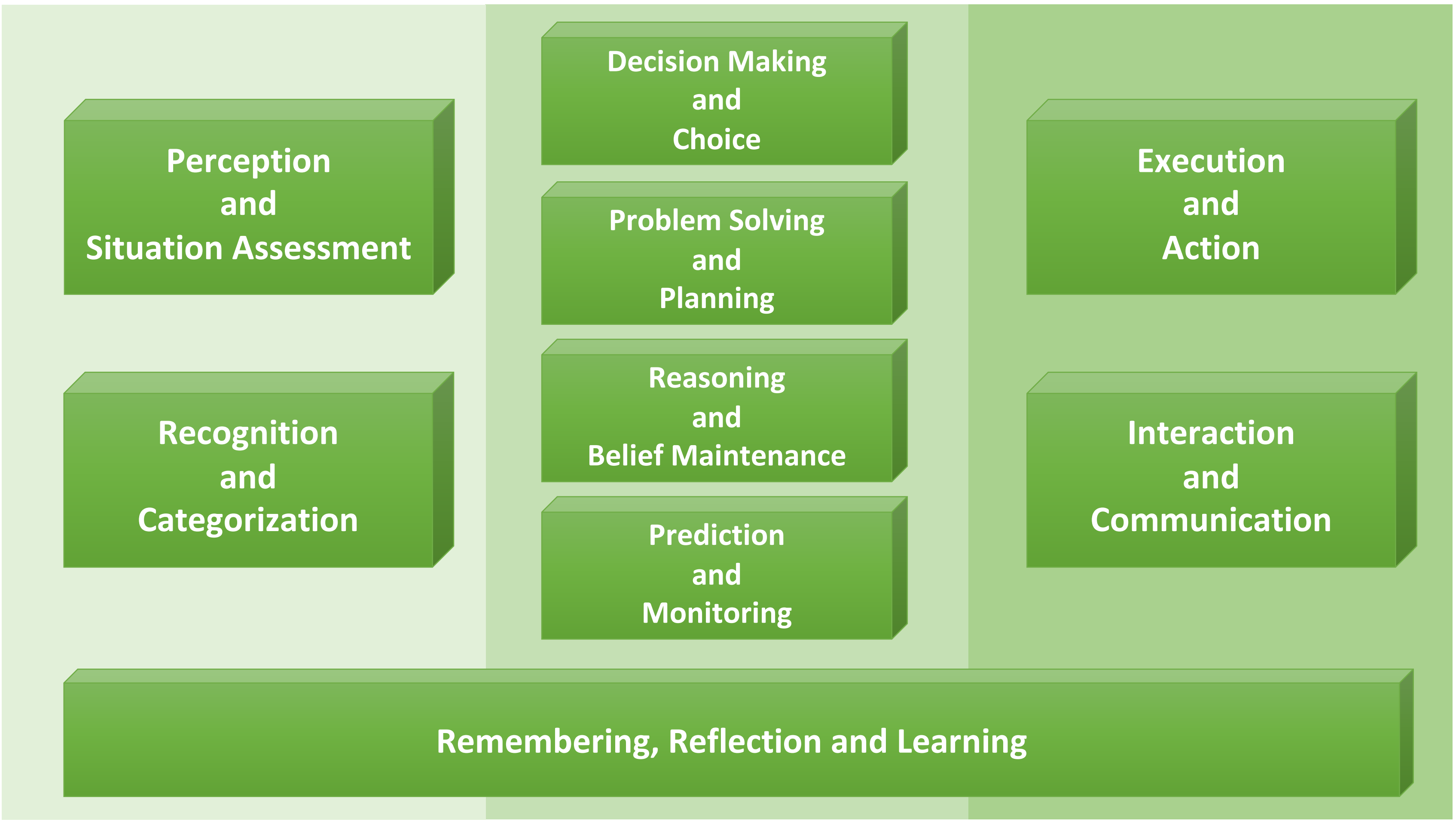
L8 – AI and ML

Topics

- The Problem Statement
- **Components of AI**
  - **Capabilities**
  - **Knowledge**
  - **Learning**
- Application Scenarios in Banking and Risk

L8 – AI and ML

Components - Capabilities



Source: *Cognitive Architectures: Research Issues and Challenges by Pat Langley, John E. Laird and Seth Rogers*

L8 – AI and ML

Components - Capabilites

Perception / Recognition / Categorization

- An intelligent agent exists in the context of the external environment that it must sense, perceive, and interpret. When using limited resources, perception will confront the issue of attention
- Sensors are often noisy and provide an inaccurate and partial picture of the agent's surroundings
- Recognition maps situations or events to known or familiar patterns. Categorization involves the assign-ment of objects, situations, and events to known concepts or categories
- A cognitive architecture must provide some way to represent patterns and situations in memory and should include some means to learn new patterns or categories from instruction or experience, or to refine existing patterns when appropriate

L8 – AI and ML

Components - Capabilites

Decision Making / Choice / Problem Solving / Planning

- A cognitive architecture must provide some way to represent alternative choices or actions and it must offer some process for selecting among these alternatives (conflict resolution, decision making)
- It should include ways to improve its decisions through learning – either by revising the conditions under which an existing action is considered allowable or by altering the numeric functions used during the conflict resolution stage
- Intelligent systems must achieve their goals in novel situations, thus they must be able to generate plans and solve problems which typically includes solution search
- To support planning, a cognitive architecture must be able to represent a plan as an (at least partially) ordered set of actions, their expected effects, and the manner in which these effects enable later actions

L8 – AI and ML

Components - Capabilites

Belief / Reasoning / Prediction / Monitoring

- The cognitive architecture must thus be able to represent beliefs and relationships among beliefs and requires mechanisms that draw inferences.Typical inferences are deductive, inductive or abductive in nature
- Reasoning draws new mental conclusions from existing beliefs or and helps deciding whether to maintain or revise those existing ones
- Prediction requires some model of the environment and the effect actions have on it. This model can also be utilized to monitor the environment against the expectations
- Monitoring relates sensing to prediction, thus it provides natural support for learning, since errors can help an agent improve its model of the environment

L8 – AI and ML

Components - Capabilites

Execution / Action / Communication

- A cognitive architecture must be able to execute skills and actions in the environment. It should also be able to learn about skills and execution policies from instruction and experience.
- For example, an agent can learn by observing another agent's behavior, by successfully achieving its goals, and from delayed rewards
- A communicating agent must represent and transform the knowledge that it aims to convey or that it believes another agent intends for it
- An important form of communication occurs in conversational dialogues, which require both generation and understanding of (natural) language, as well as coordination with the other agent in the form of turn taking.

L8 – AI and ML

Components - Capabilites

Reflection / Remembering / Learning

- A cognitive architecture can also benefit from capabilities that cut across those described in the previous slides, in that they operate on mental structures produced or utilized by them
- Remembering is the ability to encode and store the results of cognitive processing in memory and to retrieve or access them later (episodic memory)
- Reflection involves processing of either recent mental structures that are still available or older structures that the agent must retrieve from its episodic store. One type concerns the justification or explanation of an agent's inferences, plans, decisions, or actions in terms of cognitive steps that led to them
- Learning usually involves generalization beyond specific beliefs and events and involves alteration of knowledge and criteria for decision, planning or searching

L8 – AI and ML

Components - Knowledge

Representation of Knowledge

- Knowledge itself is not built into an architecture - It can change across domains and over time.
- However, the representational formalism in which an agent encodes its knowledge constitutes a central aspect of a cognitive architecture.
- One common tradition distinguishes *declarative* from *procedural* representations.
  - declarative encodings of knowledge can be manipulated by cognitive mechanisms independent of their content
  - procedural formalisms encode knowledge about how to accomplish some task.
- An architecture may also support meta-knowledge about the agent's own capabilities. Such higher-level knowledge can support meta-reasoning, let the agent 'know what it knows', and provide a natural way to an understanding of the cognitive steps taken during the agent's activities and the reasons for them
- Representational formalisms contain concepts like *Semantic Networks*, *First-Order Logic*, *Production Systems*, *Frames*, *Plans*, *Neural Networks*
  - *They differ in the way information is encoded and stored – and consequently how it can be retrieved, utilized and changed*

L8 – AI and ML

Components - Knowledge

Organisation of Knowledge

- Organisation of Knowledge influences the manner in which a cognitive architecture organizes knowledge in its memory.
- One choice that arises here is whether the underlying knowledge representation scheme directly supports *flat* or *hierarchical* structures.
  - in flat frameworks, the stored memory elements make no direct reference to each other. This does not mean they cannot influence one another.
  - In contrast, stored elements in structured frameworks make direct reference to other elements. These could be
    - a task hierarchy,
    - a part-of hierarchy,
    - an is-a hierarchy.
- Most architectures commit to either a flat or structured scheme, but task, part-of, and is-a hierarchies are complemen-tary rather than mutually exclusive
- Also, architectures need to decide on aspects of
  - the granularity of the knowledge stored in memory (fine grained units vs. plans and macro-operators)
  - the number of distinct memories that an architecture supports and their relations to each other (long term vs. short term)

L8 – AI and ML

Components - Knowledge

Utilization of Knowledge

- Utilization of knowledge stored in long-term memories can range from low-level activities like recognition and decision making to high-level ones like communication and reflection
- The desired strategy around utilization of knowledge triggers design decisions for cognitive architectures
- Given that a cognitive architecture has resource limitations which require selection among alternative goals, rules, or other knowledge structures, it needs some way to make this selection (conflict resolution)
  - Select one or more matched rules to apply based on criteria like the recency of their matched elements, the rules' specificities, or their strength
  - select moves with some numeric evaluation function that combines features of predicted states
  - systems that incorporate analogical or case-based reasoning typically select structures that are most similar to some target.

L8 – AI and ML

Components - Knowledge

Acquisition and Refinement of Knowledge

- Cognitive architectures commonly distinguish between processes that learn entirely new knowledge structures, such as production rules or plans, and ones that fine tune existing structures, e.g. through adjusting weights or numeric functions
- Learning can be done in supervised or unsupervised manner
  - Supervised learning presents known results together with a recognizable pattern and triggers the (at least partial) inclusion of this experience into the knowledge base
  - Unsupervised learning presents patterns without external classification and requires the system to self-organize this knowledge into existing knowledge categories
- Learning occurs in an incremental or non-incremental manner
  - Incremental methods incorporate training cases one at a time, with limited memory for previous cases, and update their knowledge bases after processing each experience
  - Non-incremental methods process all training cases in a single step that operates in a batch procedure

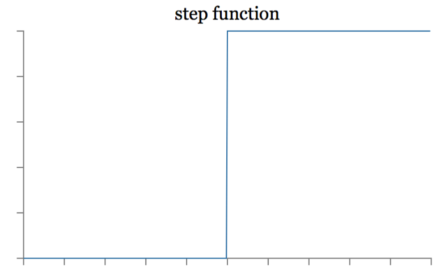
L8 – AI and ML

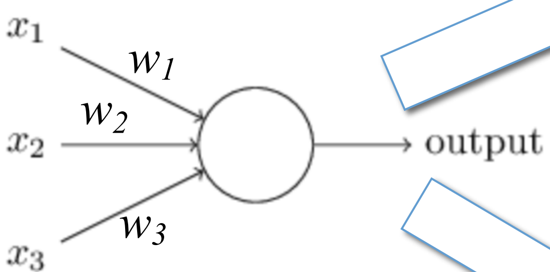
Components - Learning

Neural Networks in a nutshell:

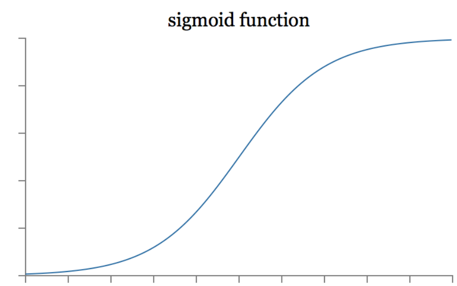
$$\text{output} = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{ threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{ threshold} \end{cases}$$

(the ‘Perceptron’)





$$\text{output} = \frac{1}{1 + \exp(-\sum_j w_j x_j - b)}$$

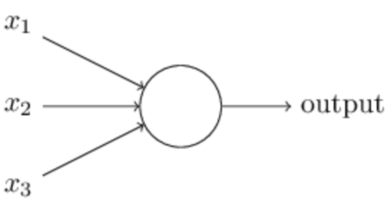


different inputs are given different ‘weights’ w

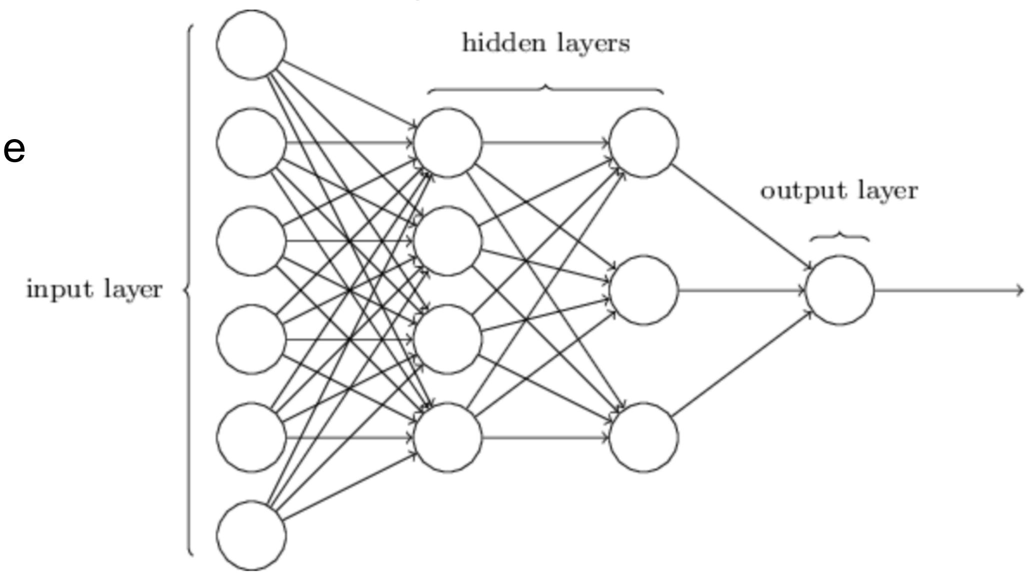
L8 – AI and ML

Components - Learning

Neural Networks can consist of many layers:



the single neuron-model can be scaled into multiple layers of multiple neurons:

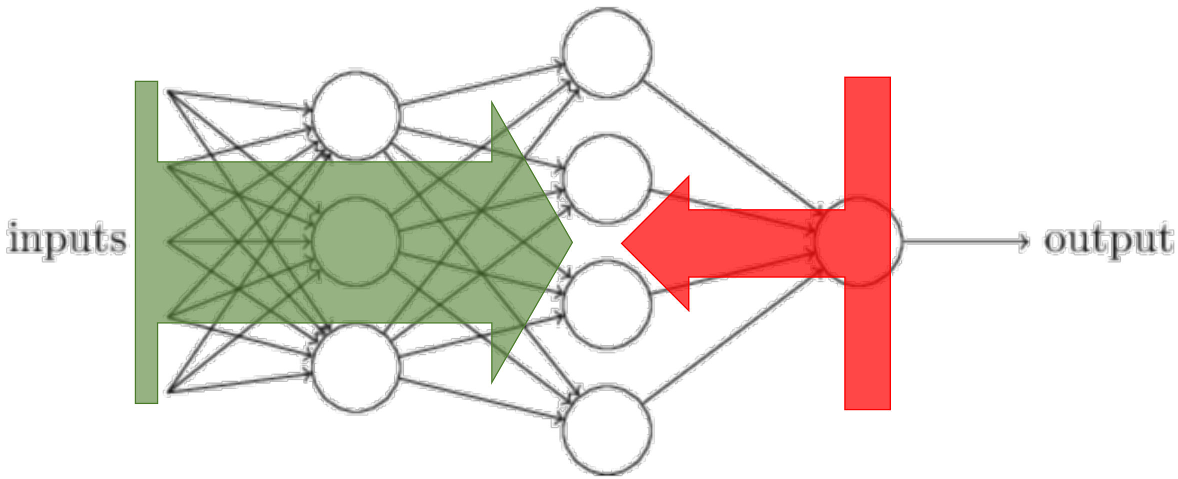


L8 – AI and ML

Components - Learning

The two directions of Neural Networks:

Learning in Neural Networks:



1) Feed forward:
   - apply input and calculate activation of hidden layers
   - calculate output (=activation of neurons in output layer)
2) Backpropagation
   - compare output from network with desired output
   - go layer by layer backwards and adjust weights to improve the distance to the target output
   - distance to target is formulated as a cost function. Minimizing the cost function happens via gradient-descent.

L8 – AI and ML

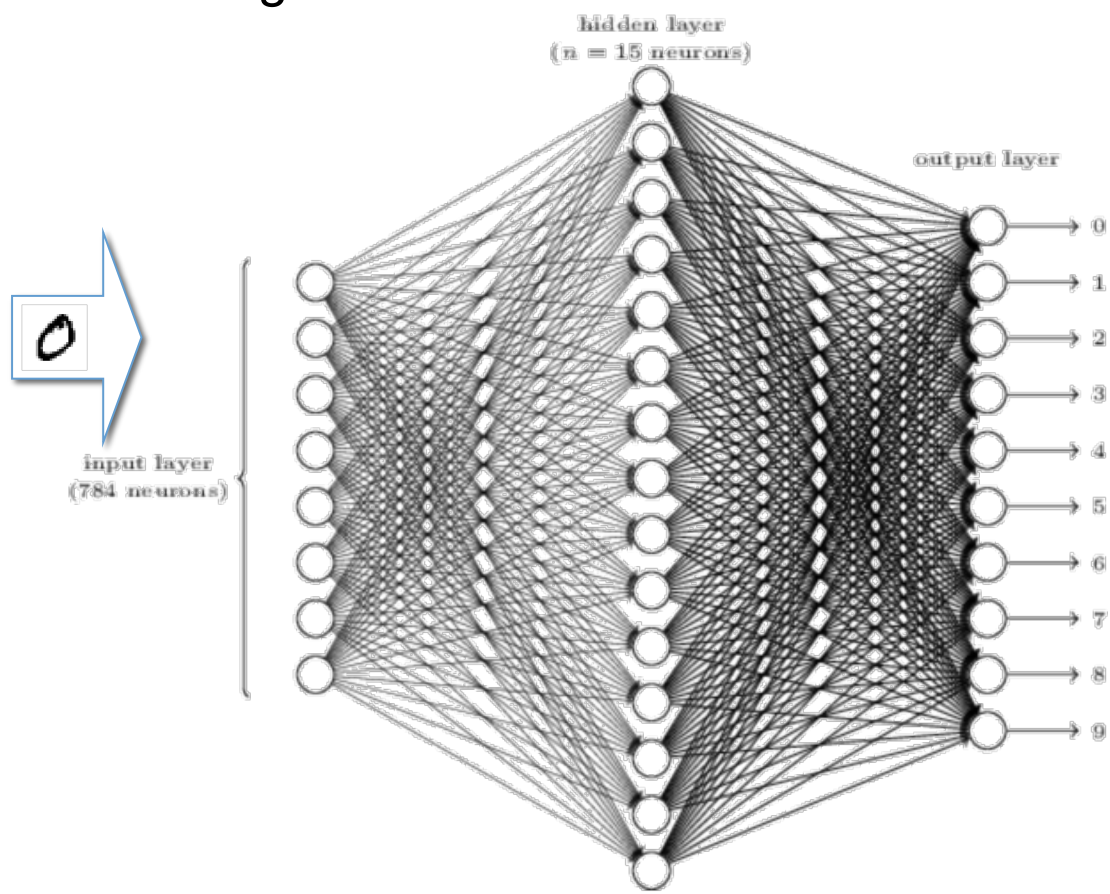
Components - Learning

Application of Neural Network Learning

Example: Recognition of handwritten digits

MNIST Dataset:

28x28 pixel per picture
= 784 pixel



L8 – AI and ML

Components - Learning

Example: Application for handwritten digits

- MNIST Dataset contains 60,000 handwritten digits
  - 10,000 of these are the test set; 50,000 are used for training
- Training can be done in multiple ‘epochs’ by
  - either using all 50,000 from the training set in random order for a single epoch
  - or using randomly chosen ‘mini-batches’ of, say, 100 out of the 50k.
- Classification accuracy is verified with the test set.
- For a simple network, the classification result can develop like this:
  - 784 input, 30 hidden neurons, mini-batch size 30 (samples per epoch)

```
Epoch 0: 9129 / 10000
Epoch 1: 9295 / 10000
Epoch 2: 9348 / 10000
...
Epoch 27: 9528 / 10000
Epoch 28: 9542 / 10000
Epoch 29: 9534 / 10000
```

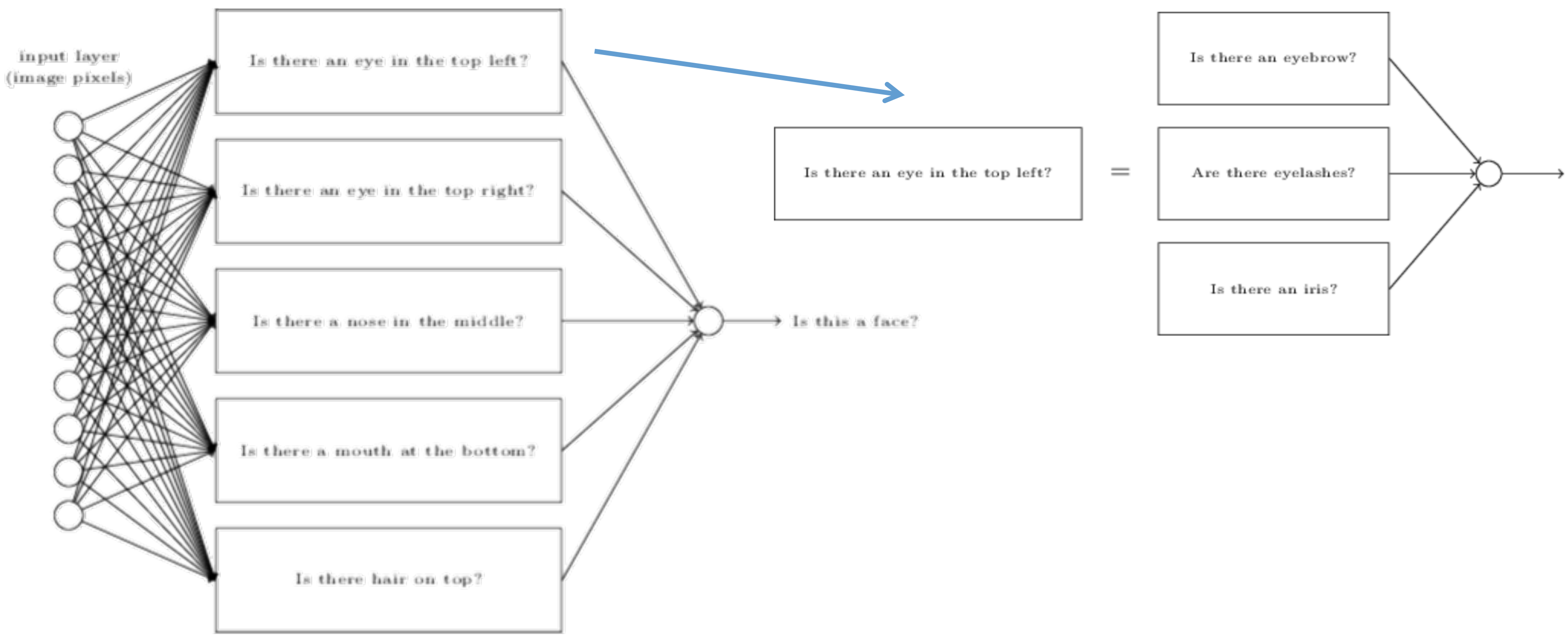
- Speed of Learning and achievable results depend on
  - structure and size of the network (=the number of free parameters)
  - Learning rate
  - ‘Representativeness’ of the training samples
  - Type and arrangement of training epochs
  - ………….

L8 – AI and ML

Components - Learning

Deep Learning is a form of Neural Network Learning

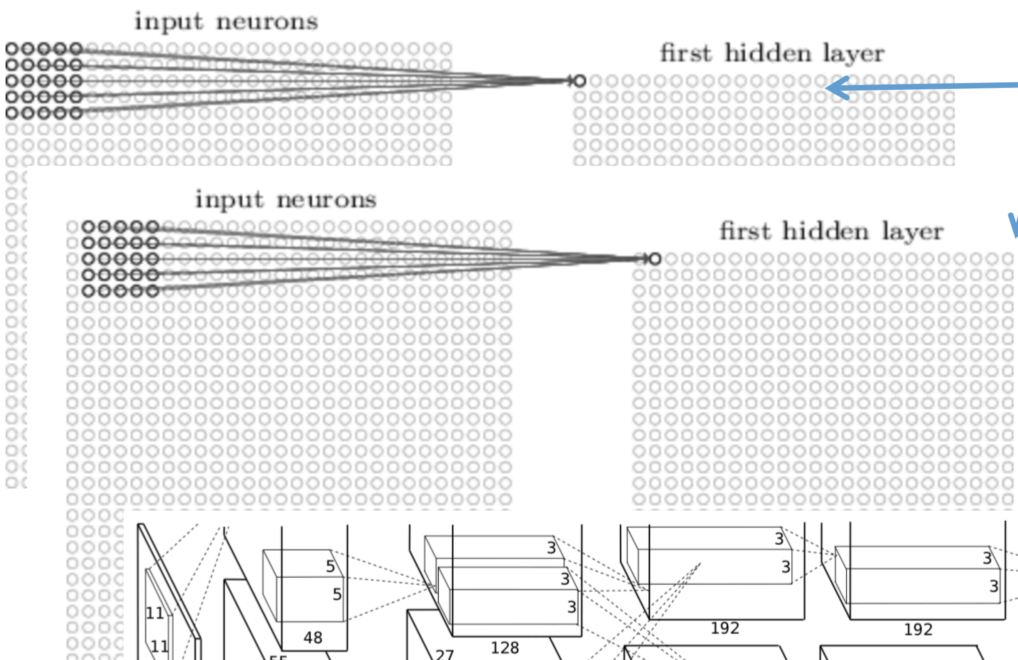
Deep Neural Networks have even more structure:



L8 – AI and ML

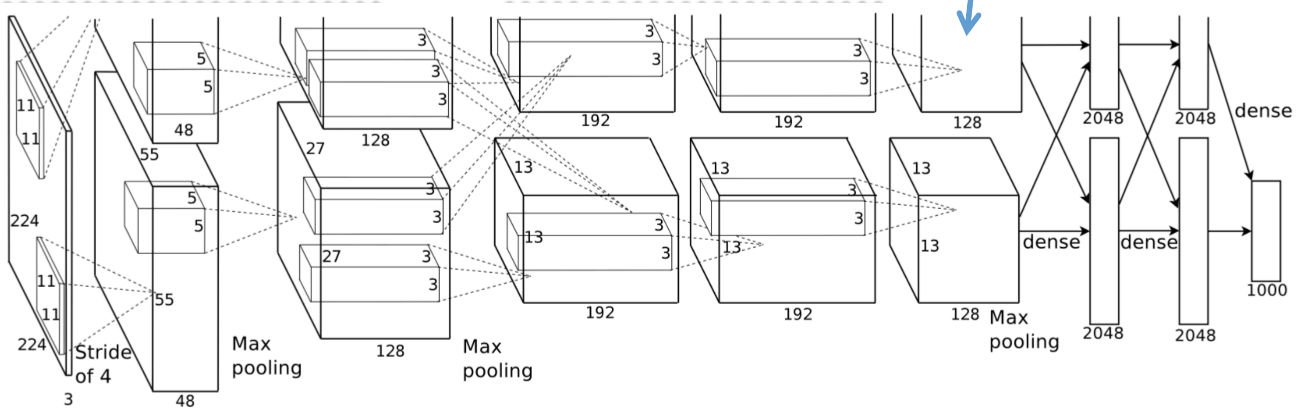
Components - Learning

Deep Learning can make use of 'feature maps'



receptive fields are created on the input layer and
are able to learn 'feature maps' in the hidden layer

depending on the task, several of such layers can be
combined to represent the target problem structure



Deep Learning involves many hidden layers to be able to extract the required features in the data

L8 – AI and ML

Components - Learning

What makes a good neural network?

- structure, composition
- good training data
- good selection of training parameters
- lots of experimentation …..

….. Neural Network design and application is not a straight forward approach !

.... Neural Networks are not the ‘ultimate weapon’ for all problems !

L8 – AI and ML

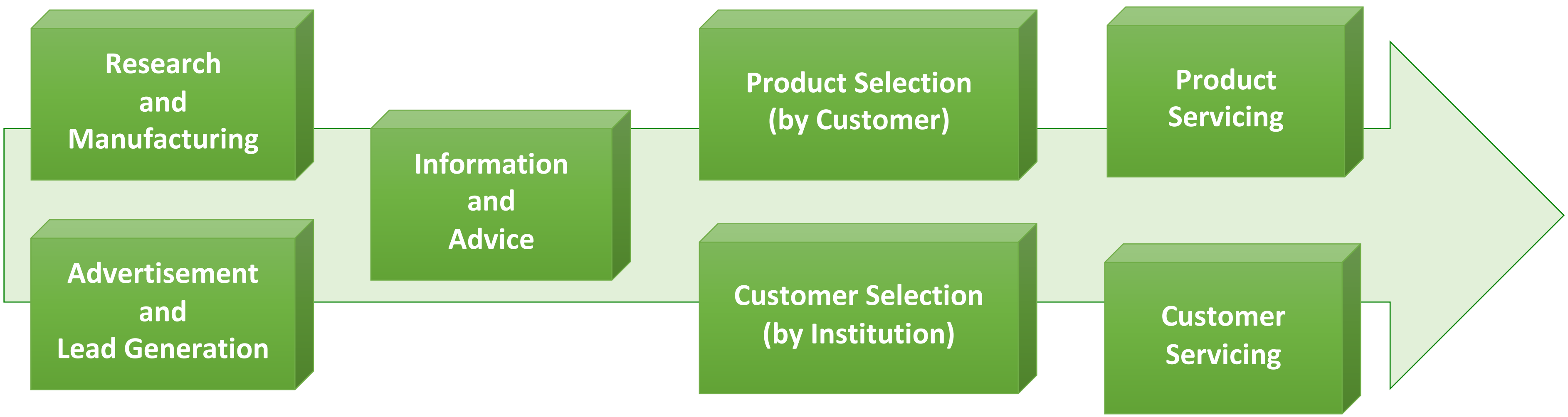
Topics

- The Problem Statement
- Components of AI
- **Application Scenarios in Banking and Risk**

L8 – AI and ML

Application scenarios in Banking and Risk

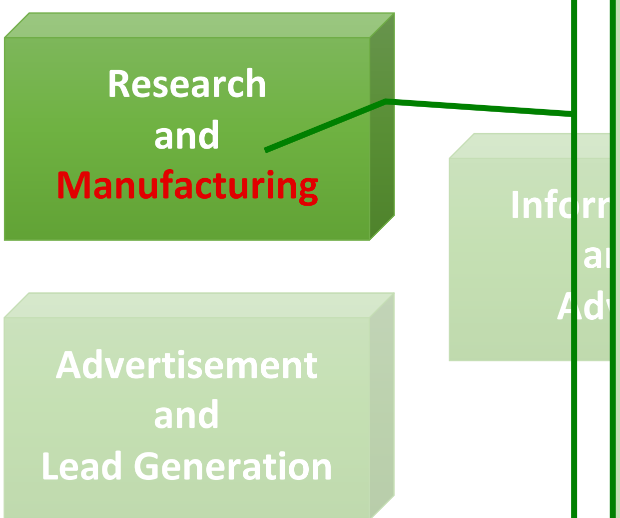
- **A generic business process for financial products**



L8 – AI and ML

Application scenarios in Banking and Risk

- **A generic business process for financial products**



Sample Objective:	Comply to Rules and Regulations (KYC, AML, national specifics, …)
Application of AI:	Regularly crawl regulator's announcements Natural / legal language processing Recognition / Categorization Autonomous forming of a knowledge base
Benefits of AI:	Augment human research capabilities Extend speed and reach Prepare for automated compliance checks
Points of Caution:	Requires human vetting of results

L8 – AI and ML

Application scenarios in Banking and Risk

- **A generic business process for financial products**

Research and Manufacturing

Advertisement and Lead Generation

Information and Advice

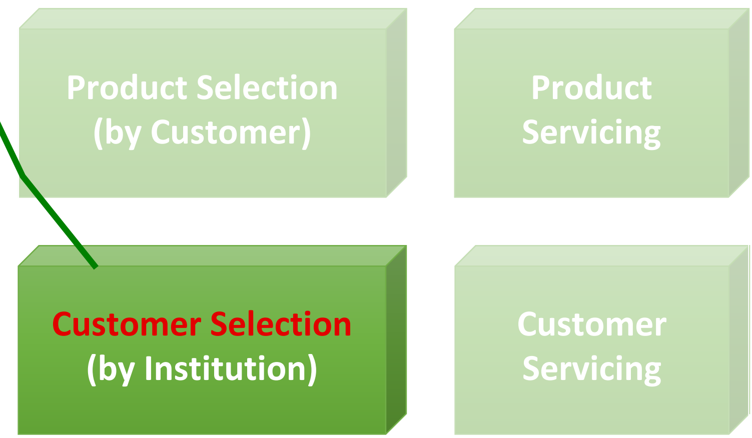
Sample Objective:	Answer queries of potential customers Engage and inform about products
Application of AI:	Chat Bot for natural language conversation Sentiment Analysis (conversation patterns) Communication of product knowledge
Benefits of AI:	Automate simple conversations Reduce 'tedious' human resource binding Ensure consistence of information given out
Points of Caution:	Requires human 'back-out plan' Online/incremental learning dangerous

L8 – AI and ML

Application scenarios in Banking and Risk

- **A generic business process for financial products**

Sample Objective:	Minimize Customer related risk Third Party Due Diligence, KYC Credit Risk evaluation
Application of AI:	Natural Language processing Pattern Recognition Prediction Decision Making
Benefits of AI:	Broaden scope of investigations Faster decisions
Points of caution:	Apply '2I' principle for bigger risks Learn from false positives/negatives



2 Intelligences: 1 artificial, 1 human

L8 – AI and ML

Application scenarios in Banking and Risk

- **A generic business process for financial products**

Sample Objective:	Monitor Customer related risk Verify own prediction Ongoing Credit Risk evaluation
Application of AI:	Pattern recognition on payments Monitoring, Prediction
Benefits of AI:	Broaden scope of investigations Automated monitoring Improved prediction capability
Points of Caution:	Apply ‘2I’ principle for bigger risks Learn from false positives/negatives

Product Servicing

Customer Servicing

L8 – AI and ML

Application scenarios in Banking and Risk

The risk assessment of cognitive architectures themselves is a challenge. Aside from the common risks of complex systems, there are new risks and questions coming up, e.g.:

- A cognitive architecture is aimed to perform 'tasks not explicitly programmed for'
  - This gives rise to new challenges in Quality assurance and testing
  - How can such a system be 'certified' against requirements and design?
- A 'creative' cognitive architecture is difficult to explain
  - Can we understand exactly what the system did and why?
  - How do we explain to regulators what the system did?
- Can we allow incremental learning (online Learning)?
  - incremental learning happens as the systems works
  - but how do we control it is learning the 'right things'? What are the 'right things'?
- Who is ultimately accountable?
  - Conventional thinking of 'education', 'qualification' etc. does not hold in this context
  - Thus, how to assess the capability? How to assign responsibility? What is the consequence of accountability?

L8 – AI and ML

Application scenarios in Banking and Risk

Guiding Principles for thinking about Intelligent Systems and Risk Management:

- Artificial Intelligence is not a uniform algorithm or solution space. It consists of a variety of capabilities, different forms knowledge and abilities to adapt.
- Machine Learning is a 'subordinate' concept, attempting to improve the existing knowledge base of a cognitive architecture.
- Practical approaches should view AI as a toolbox and choose the solutions required based on the problem space and the underlying risk model
- Artificial Intelligence – properly designed and applied - helps to identify, assess, monitor and mitigate risk in financial products and processes as well as to automate recurring processes
- Risk Management needs to develop a completely new approach to cover the additional risks introduced by Artificial Intelligence

Use AI – but choose wisely