Neural Networks

Introduction

• **Neural:** $\Longrightarrow$. Neuron. Basic unit (cell) of nervous systems.

• **Network:** $\Longrightarrow$. An arbitrary interconnection of predetermined blocks with well defined input/output relations.

• Arbitrary interconnection only makes sense from mathematical point of view. From engineering point of view, the connections should be such that the network should perform a predetermined task .

- Computer Networks.
- Communication Networks.
- Power System Networks, etc.

• Basic building block (unit) of a Neural Network is a neuron. A neuron is a very complex biological unit, its functioning is not completely understood yet. Naturally we will use well-defined mathematical relations between well-defined inputs and outputs of a neuron $\Longrightarrow$ A mathematical Model.

• In determining the mathematical model, we study the basic functioning of human neuron cell.

• **Neural Network:** Arbitrary connections of such neuron units. But we will use the well-defined mathematical relations of a neuron model.

- **Our Aim:** To be able to perform certain tasks that our nervous system does naturally which classical control/information systems can not solve easily.

    - Learning.
    - Pattern Recognition.
    - Survival in an environment.
    - Motion Planning

- **Neural Network Theory:** Any study which investigates any aspect of neural networks .

- The subject is naturally highly interdisciplinary. $\Longrightarrow$ Biology, Physics, Chemistry, Almost all branches of Engineering, Cognitive Sciences, etc...

- Engineering Approach :

    - Develop the basic mathematical models of basic units .
    - Analyze the behaviour of appropriate interconnections of basic units.
    - Find appropriate architectures based on models.
    - Apply these models and develop methods to solve certain engineering problems.

- We will focus on :

    - Learning.
    - Pattern Recognition.

- **Basic Orientations in Neural Network Theory :**

- **Cognitive Sciences/Artificial Intelligence :** Modelling intelligent behaviour, to overcome the classical symbolic methods of cognitive science, i.e. to analyze these concepts from mathematical point of view.

- **Neurobiological Modelling:** To model the behaviour of individual neuron. Exact properties of neuron plays important role.

- **Scientific Modelling:** To use the Neural Networks as basic modelling tool for various systems in e.g. physics, psychology, sociology, etc.

- **Computer Science :** Views Neural Networks as a special class of algorithms that has special properties (e.g. fault-tolerant, generalization, etc.)

- **Engineering Applications :**
    - Classification. Any problem which can be recast as a classification problem is a potential candidate for Neural Networks application.
    - Pattern Recognition. Can be considered as a special case of classification.
    - Optimization. Neural Networks can be applied to a wide class of optimization problems.
    - Control. Many complex control problems can be solved by using neural techniques.
    - Signal Processing. Many problems which are related to classification and pattern recognition which are classically related to signal processing can also be solved by neural techniques (e.g. speech recognition...).

- **Types of Neurons :**

  - Interneuron. From Neurons to Neurons. Majority of neurons belongs to this class.

  - Sensory Neurons. connected to the sensory organs, e.g. eyes, ears, skin, etc..

  - Motor Neurons. Connected to muscles.

- All types have different structures, sizes, shapes, but have the same main components :

  - Dendrites. Input ports. Collects “data” (potential) from nearby neurons. Cylindrical shape. Radius $\sim$ a few micron.

  - Soma. Main cell body. Collects the potential received from dendrites. Size : $\sim 10 - 80$ microns.

  - Axon. Transmission line. Cylindrical shape. Radius : $\sim 0.2\mu - 1mm$. Length : $\sim$ a few μ- 1 m!. Long axons are connected to motor neurons and are covered by myelin sheet for insulation which increases the speed.

  - Synapses. Transmits the data to the next neuron dendrites. Cylindrical shape. Radius $\sim$ a few micron.

  - Total Length : 10μ (for interneurons) - 1 m (for motor neurons).

  - **Information Flow :** One direction only. (Neuron 1 : Dendrites $\rightarrow$ Soma $\rightarrow$ Axon $\rightarrow$ Synapses $\rightarrow$ (Neuron 2 : Dendrites .....).

  - Synaptic Gap. Between a synapse and dendrite. $\sim$ 200 nm.

- **Signal Generation :**

- **Neutral Situation :** No signal from outside. Behaves like a battery. Basic agents are the sodium (Na) and potassium (K) ions which are positive. Less important agents are chloride (Cl) ions, which are negative and large organic ions.

- Neutrally, cell membrane has more resistance (less conductance/permeability) to Na ions as compared to K ions (ratio ~ 1000). As a result, there are more K ions and less Na ions inside as compared to outside. Similarly, there are more Cl ions outside as compared to inside. Organic ions are too large to pass the cell membrane too often.

- **Result :** A chemical balance occurs. Around $70 - 100\ mV$ potential occurs between inside and outside of the cell, inside being more negative. $-70\ mV$ is more common. This is called **Resting Potential**.

- **Excitant Situation :** Soma receives potentials through its dendrites which excite them through the synapses of the connected neurons. The effect of this decays in $5 - 10\ msec$. This is called **latent summation**. This changes the potential inside the soma. When a certain threshold is passed (typically $-60\ mV$), the resistance of cell membrane wrt. Na ions drops **drastically**, while resistance to K ions remains the same for some period. The net result is a sudden influx of Na ions inside the cell, which increases the cell potential (typically to $+30\ mV$). This is called **Action Potential**. During this period, the cell stops receiving any input from is dendrites for $1 - 5\ msec$ (**resting period**). Then the resistance of cell membrane wrt. Na and K ions returns to their normal values and the resting potential is restored. Then the cell undergoes another (**refractory**) period, a few msec., during which no input to the cell is allowed.

• This process is discovered and modeled mathematically by Alan Hodgkin, Andrew Huxley and John Eccles. They received 1963 Nobel Price for this discovery.

• **Signal Transmission :**

• Action potential travels along the axon and reaches synapses almost without attenuation for short axons (interneurons). For long neurons (motor neurons), to achieve this, axon is covered by insulating sheets called myelin. depending on the length, there may be a number of such sheets, separated by what is called **Ranvier Nodes**. At these nodes, the same action potential again generated by the same principle, and reaches the synapses almost without attenuation.

• When action potential reaches the synapses, they release chemical substances what is called **neurotransmitters** (e.g. 10^{-17} mol. acetylcholin per impulse). These neurotransmitters reaches the next dendrite around 0.5 *msec*. Their role is to change the permeability of the other side with respect to Na, K ions (to increase the potential → **excitory synapse**), or to Cl ions (to decrease the potential → **inhibitory synapse**). This behaviour is fixed in the early stages of life after some learning process. But the synapses may have different **strengths** which affect the amount of released neurotransmitters.

• This process is discovered by Henry Dale and Otto Loewi. They received 1936 Nobel Price for this discovery.

- **The speed of transmission :** $0.5 - 2\ m/sec$ in interneurons.

- Any two brain cell can communicate in about $20 - 40\ msec$.

- In long axons, the speed of transmission can go as high as $100\ m/sec$.

- The frequency of spikes : $\leq 1\ KHz$. (signal duration $\sim 1\ msec$).

- **Is the information processed digitally ?** Not exactly. There are evidence that the information may also be coded in the frequency of the signal.

- **How many neurons are there?**

  $\sim 3 \times 10^{10}$ in the brain, $\sim 10^{11}$ in the nervous system.

- Number of synaptic connections changes from neurons to neurons. On the average, a neuron makes $\sim 10^4$ synaptic connections. $\rightarrow \sim 10^{15}$ total connections !

- $\rightarrow$ We may compare the brain (or nervous system) with a computer with $\sim 10^{11}$ processors with $\sim 10^{15}$ connections! Impossible to realize even in the near future!

- Energetic efficiency of computers : 10^{-6} Joules per operation per second.

- Energetic efficiency of neurons : 10^{-16} Joules per operation per second! Highly efficient.

• Speed of computers : $\sim 10^{-9}$ second per operation. Very fast and even decreasing with newer technology.

• Speed of neurons : $\sim 10^{-3}$ second per operation. Unfortunately impossible to change!

• Classical computers are useful in problems whose solutions have well defined algorithms which can be implemented serially. → multiplication, linear programming, etc...

• Neural Networks are more efficient in cognitive related tasks which does not have well defined algorithms. → pattern recognition, problem solving, motion coordination and planning etc..

• **How does the neural connections form?**

• A minority of these connections comes with birth. But the majority are formed afterwards. Most of them occur in the first 2 years of life. During this period, $\sim 10^6$ connections are formed per second! This is called **hard wiring of the brain**. Development continues at later ages but slows down. During this period, connections may be made, later deleted, strengths of synapses modified etc. This is called **LEARNING**.

• **How does the brain decide on connections ?** → We do not know. We concentrate on some form of optimization. Mainly, because of engineering/physical background this is the first think which comes to our mind.

• **Is the structure genetically coded ?** → Some part may be. For example, ability to see is inborn. But speech is not. One has to learn how to speak with the help of the environment. Even some properties related to vision (e.g. pattern recognition) must be learned. → Appropriate synaptic connections must be formed. → Kitten experiment..

• **Why to study Neural Networks ?**

• Historically. the interest in Neural Networks has two roots :

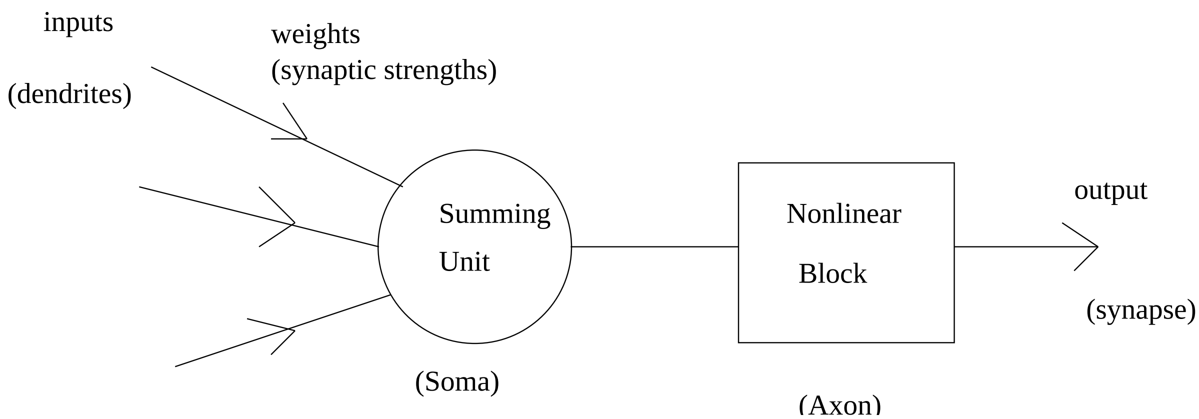
• To understand how our brain and nervous system works.

• To build machines/develop algorithms that are capable of doing complex tasks which are not easily solved by classical computing techniques.

• Recent interest in engineering is focused on the second approach.

• **One major disadvantage :** We still do not know how our brain solves such complex tasks. To be able to solve them, we still think in terms of the methods we know in engineering (such as optimization)

- **A simple Neuron Model :**



- **A Neural Network:** An arbitrary connection of such units to perform a given task.

- **Problem :** Find appropriate structure and weights to solve the task → Learning.

- **Benefits of Using Neural Networks:**

- **Nonlinearity :** Makes it useful in certain applications such as function approximation, optimization...

- **Input/Output Mapping :** Neural Networks can approximate a given input/output mapping. Useful in control design, function approximation...

- **Adaptivity :** Weights can be changed. Neural Networks may adapt themselves to changing environmental conditions to give better response.

- **Fault Tolerance :** Due to a large number of interconnection, even a part fails, the system response may not degrade drastically.

- **VLSI Implementability :** Basic unit can be realized electronically, and neural network structure can be realized in VLSI chips.

- **Neurobiological Analogy :** By better understanding basic neurons, we may develop better models and based on these models we may better understand the brain functioning → role of dreams....

- **Examples of simple tasks that can be done by Neural Networks**

- **Simple Classification :**

- Given points on the hypercube $P_0 = (-1, -1, -1),\ P_1 = (-1, -1,\ 1), \ldots$
- $P_7 = (1, 1, 1)$

- **Classification Problem :**

- $$P_i \in \begin{cases} \mathcal{C}_1 & \#1 > \# -1 \\ \mathcal{C}_2 & \#1 < \# -1 \end{cases}$$

- **Model :** $o = f(x_1 + x_2 + x_3)$

- $$f(x) = \begin{cases} +1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

- This is called **signum** function.

- Here $x_i = \pm 1$ are the inputs, $w_i = 1$ are the weights (excitory synapses, $i = 1, 2, 3$)

- for $P_0 = (-1, -1, -1) \rightarrow x_i = -1 \rightarrow o = f(-3) = -1$

- for $P_7 = (1, 1, 1) \rightarrow x_i = 1 \rightarrow o = f(3) = 1$

- $$P_i \in \begin{cases} \mathcal{C}_1 & o = +1 \\ \mathcal{C}_2 & o = -1 \end{cases}$$

- Divides the space into two parts $\mathcal{C}_1 : x_1 + x_2 + x_3 > 0$, $\mathcal{C}_2 : x_1 + x_2 + x_3 < 0$

- Patterns are **linearly separable**

- **Function Approximation :**

- $o = F(W, x)$ is the output of the neural network. F is the function of the nonlinear block (and fixed), x is the input (given), and W is the set of the weights (to be determined).

- **Problem : Given a function** $y = f(x)$, $x \in \Omega$, find appropriate weights W such that $\mid o - y \mid$ is as small as desired over the set $x \in \Omega$.

- Any continuous function can be approximated by a neural network with any desired accuracy.

- Has many applications, e.g. in control theory. Any system with given input/output relations (or experimental results) can be approximated. $\rightarrow$ Generalization....

• **Simple Memory and Restoration of Patterns :**

• Need differential or difference equations.....

• $o(k+1) = F(W, o(k)) \qquad k = 0, 1, 2, 3, \ldots$

• Here $o(0)$ is the initial pattern, $o(k)$ is the pattern at step k, etc...

• For o^* to be a memory pattern, we should have $o^* = F(W, o^*) \rightarrow$ fixed point....

• For $o(0) = o^* + \Delta$ be a distorted version of o^*. If we can have $o(k) \rightarrow o^*$, then we can restore the pattern o^* from its distorted version. $\rightarrow$ stability

• **Brief History :**

• Human central nervous system had been studied by medieval doctors since the Middle Ages...

• In the second half of 19th century, there were two theories :

• **i :** Nervous system consists of a continuous (uninterrupted) network of nerve fibres.

• **ii :** Nervous system consists of an interconnection of vast amount of simple nerve cells.

• Towards the end of the century, a chemist (Golgi) developed a technique to detect the tiny gaps in matters. By using this technique, a Spanish doctor Ramon Cajal proved in 1888 that there are tiny gaps in nerves, proving the second theory. (He detected the gap between synapse-dendrite pairs, which is ~ 200 nm). They received the 1906 Nobel price..

• Around 1911, various neuron models were developed, some by Cajal.

• After the invention of electron microscope, various types of neurons and their structures were identified. (1950).

• Modern era of neural networks started with the works of McCulloch (neurobiologist) and Pitts (mathematician). An elementary mathematical model of computing neuron. Can perform basic logic operations. Similar to Arithmetic Logic Unit of classical computer.

• McCulloch-Pitts model raised a lot of interest. J. Von Neumann (famous mathematician, builder of first computer) gave lectures on this model. His works are collected in a book : The Computer and the Brain (1958).

• N. Wiener (famous mathematician) wrote a book called : Cybernetics in 1948. He tried to relate statistical mechanics with learning and self-organization. The same linkage was considered in 1980s by other scientists.

• In 1949, D. Hebb published a book : The organization of Behaviour. He proposed a learning technique called Hebb's Rule.

• In 1954, Minsky wrote his PhD thesis on Neural Networks.

• In 1954, Gabor (inventor of holography, pioneer in communication theory, Nobel Prize winner), introduced the idea of nonlinear adaptive filter. He developed a special learning concept by using stochastic systems..His team built such a device around 1960..

- In 1958, Rosenblatt introduced a model called perceptron. Capable of classifying some class of patterns. Proved this mathematically.

- In 1960, Widrow and Hoff introduced a powerful learning rule. They build a device called ADALINE (Adaptive Linear Combiner), which is used in pattern recognition and adaptive filtering/control. Used in filtering noise in telephone lines.

- In 1969, Minsky and Papert wrote a book and demonstrated the basic limitations of single layer perceptrons. This work was influential and affected the works on Neural Netwoks adversely in 70s.

- In 1980's, Grossberg and Carpenter introduced a number of neural network models. This lead to the adaptive resonance theory.

- In 1982, Kohonen introduced his work on self organizing neural networks.

- In 1982, Hopfield (biologist) introduced a recurrent neural network model. His work was very influential.

- In 1986, Rumelhart, Hinton and Williams reported the back propagation algorithm. This introduced a powerful training rule for multilayer networks. It turned out that the sam algorithm was introduced independently by others in 1985, and even in the PhD thesis of Verbos in 1974.

- In 1989, the first VLSI neural network chips were designed.