CHAPTER III

THEORETICAL BACKGROUND

3.1 An Overview of Neural Computing

Refer to Neural Networks textbook [19], the field of artificial intelligence (AI) has made a great progress in the direction of automating human reasoning. Nevertheless, the tools of AI have mostly restricted to sequential processing and only certain representations of knowledge and logic. A different approach to intelligent system involves constructing computers with architectures and processing capabilities that mimic certain processing capabilities of the brain. The results are knowledge representations based on massive parallel processing, fast retrieval of large amounts of information, and the ability to recognize patterns based on experience. The technology that attempts to achieve these results is called neural computing or *artificial neural networks*, networks made up of neurons in much the same way as a brain. A biological neuron consists of four major components, namely, soma, axon, dendrite, and synapse. The input and output signals to the soma of a neuron are transmitted along the axon and dendrite. The synaptic resistance controls the strength of the signal. A neuron learns to generate a particular signal by adjusting the synaptic resistance. An artificial neural network is a mathematical model that emulates a biological neural network. The similarity between the biological neuron and the artificial neuron is summarized as follow.

Biological Neuron	Artificial Neuron
Soma	Node
Dendrite	Input

Axon	Output
Synapse	Weight

In neural networks, we carry over this concept of synaptic resistance and refer to it as the weight of a neuron. There are three types of learning mechanism dealing with the adjustment of a neuron weight. The first type is called *supervised learning*. A neuron is forced to generate a target signal associated with a specific input pattern and to reproduce this target signal whenever the specific input pattern occurs. The second type of learning is called *unsupervised learning*. There is no target signal generated with a particular input pattern. A neuron competitively adjusts its weighted value with the other neurons to make the value of its weight equal to the value of the input pattern. The last type of learning is called *reinforcement learning*. This type is a mixture of the first two types under the environment that the target is specified. In this research, we are concerned with only the first type of learning.

3.1.1 A Backpropagation Neural Network

Refer to textbook of Introduction to The Theory of Neural Computation [14], the Back-Propagation algorithm is vital to much work on learning in neural networks. The algorithm gives an instruction of changing the weights in any feed-forward network to learn a training set of input-output pairs $\{ \xi_k^\mu, \zeta_i^\mu \}$. We consider first a two-layer network such as that illustrated in Figure 3.1. Our notational conventions are shown in the figure; output unit are denoted by O_i, hidden units by V_i, and input terminals by ξ_k. There are connections W_{jk} from the input to the hidden units, and W_{ij} from the hidden units to the output units. Note that the index i always refers to-an output unit, j to a hidden one, and k to an input terminal. The inputs are always taken to particular values. We label different patterns by a subscript μ, so input k is set to ξ_k^μ when pattern is being presented. The ξ_k^μ can be binary or continuous-valued. We use N for the number of input units and p for the number of input patterns.

Given pattern μ, hidden unit j receives a net input

$$h_j^\mu = \sum_k W_{jk} \xi_k^\mu$$

and produces output $$V_j^\mu = g(h_j^\mu) = g(\sum_k w_{jk} \xi_k^\mu).$$

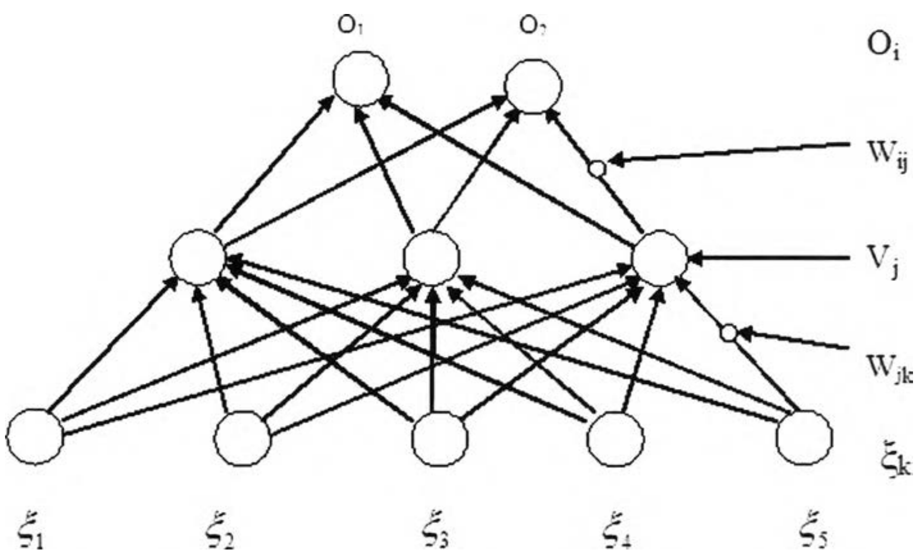

Figure 3.1: A two-layer feed forward network, showing the notation for units and weights.

Thus, output unit i receives

$$h_i^\mu = \sum_j W_{ij} V_j^\mu = \sum_j W_{ij}\ g(\sum_k w_{jk} \xi_k^\mu)$$

and produces for the final output

$$O_i^j = g(h_i^j) = g(\sum_j W_{ij} V_j^\mu) = g(\sum_j W_{ij} \; g(\sum_k w_{jk} \xi_k^\mu)).$$

Our usual error measure or ***cost function***

$$E(W) = \frac{1}{2} \sum_{\mu i} [\zeta_i^\mu - O_i^\mu]^2$$

now becomes

$$E[W] = \frac{1}{2} \sum_{\mu i} \left[\zeta_i^\mu - g(\sum_j W_{ij} \; g(\sum_k w_{jk} \xi_k^\mu)) \right]^2 .$$

This is clearly a continuous differentiable function of every weight. So we can use a gradient descent algorithm to learn appropriate weights.

For the hidden-to-output connections, the gradient descent rule gives

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} = \eta \sum_\mu [\zeta_i^\mu - O_i^\mu] g'(h_i^\mu) \, V_j^\mu$$

$$= \eta \sum_\mu \delta_i^\mu V_j^\mu$$

where $$\delta_i^\mu = g'(h_i^\mu)[\zeta_i^\mu - O_i^\mu]. \qquad (1)$$

For the input-to-hidden connections Δw_{jk} we must differentiate with respect to the w_{jk}'s.

Using the chain rule, we obtain

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \sum_\mu \frac{\partial E}{\partial V_j^\mu} \frac{\partial V_j^\mu}{\partial w_{jk}}$$

$$= \eta \sum_{\mu i} [\zeta_i^\mu - O_i^\mu] g'(h_i^\mu) \, W_{ij} \; g'(h_j^\mu) \, \xi_k^\mu$$

$$= \eta \sum_{\mu i} \delta_i^\mu W_{ij} \; g'(h_j^\mu) \xi_k^\mu$$

$$= \eta \sum_\mu \delta_j^\mu \; \xi_k^\mu$$

with $$\delta_j^\mu = g'(h_j^\mu)\sum_i W_{ij}\delta_i^\mu \, . \qquad (2)$$

In general the backpropagation update rule always has the form

$$\Delta w_{pq} = \eta \sum \delta_{output} \times V_{input}$$

where *output* and *input* refer to the two ends p and q of the connection concerned, and V stands for the appropriate input-end activation from a hidden unit or a real input. The meaning of δ depends on the layer concerned. For the last layer of connections, it is given by an equation (1), while for all other layers it is given by an equation (2). Equation (2) allows us to determine δ for a given unit V_j in terms of the units O_i that it feeds. The coefficients are just the usual "forward" W_{ij}'s, but, here, they are propagating errors $(\delta$'s$)$ backwards instead of signals forwards.

3.2 Weather Radar Fundamentals

Reference of Satellite Meteorology: An Introduction [15], transmitting a pulse of microwave radiation and making measurements of the radiation returned from precipitation or other objects are performed by radar which measures the *power*, *frequency*, and *phase* of the returned signal. The returned power (P_r), which depends on properties of a particular radar (antenna area, wavelength, transmitted power, pulse duration) and $Z = \sum_{vol} D^6$ are obtained. $Z = \sum_{vol} D^6$, called *radar reflectivity factor* (Z), is usually measured in units of millimeters to the sixth power per cubic meter (mm^6 / m^3) and D means diameter of droplet. In decibels; $dB = 10 \log_{10}(P_r / P_{ref})$, the power returned to a radar is measured. For the Rayleigh drops, the reference power P_{ref} is often taken to be that power which would be returned if each cubic meter of atmosphere is filled with one drop with diameter of 1 mm. ($Z = 1\ mm^6 / m^3$). The returned power is indicated by the symbol dBZ, when referenced in this way. The Mie scattering equations can be used to calculate the power returned to the radar, if the droplets are not in the Rayleigh region. In contrast, an *effective radar reflectivity factor* (Z_e) is calculated. In this case, if one has an estimate of the drop size distribution and the velocity at which water drops of a particular diameter fall, one can use the radar reflectivity factor to calculate the precipitation rate. These relationships often take the form $Z = a\ R^b$, which is known as a $Z_e - R$ relationship. The rainfall rate R is measured in units of millimeters per hour (mm/hr).

Horizontal motions of drops in the sample volume cause frequency shifts in the returned signal due to the Doppler Effect. There are two types of frequency shifts, which are random motion of the drops causes a spread, but no shift, in the spectrum of returned radiation, and another one is systematic frequency shifts are caused by non-random translation of the drops. If the frequency of the returned radiation is measured with a *Doppler radar*, the speed of the translation of the drops in the radial direction (toward or away from the radar) can be calculated. But even these small shifts can be accurately measured to yield radial drop speed and, thus, an estimation of the radial wind speed. NEXRAD (next-generation radar), is a Doppler radar.

The phase of the returned radiation can be measured. Information on drop shape can be obtained by comparing the phases of horizontally and vertically polarized measurements. Since large drops associated with higher rainfall rates tend to be flatten as they fall, the difference in phase between the two polarizations can be used to estimate rain rate.

3.2.1 Mie and Rayleigh scattering

Reference in [15], Radiation scattered from a particle is a function of several things: particle shape, particle size, particle index of refraction, wavelength of radiation, and viewing geometry. For size parameters in the range 0.1-50 of Mie scattering, the wavelength of the radiation and the circumference of the particle are comparable. When radiation strongly interacts with the particle, the full Mie equations, which have been applied extensively to the detection of raindrops by radar, must be used. The study of aerosols (smoke, dust, haze) using visible radiation falls in the Mie regime, and also in the Mie regime is the interaction of cloud droplets with infrared radiation. In contrast, size particles less than about 0.1, Rayleigh scattering is insensitive to particle shape; works for nonspherical as well as spherical particles. Also, air molecules act as Rayleigh scattering for visible and ultraviolet radiation. This is why Z and Z_e is different.

3.3 Convective and Stratiform Precipitation

3.3.1 Precipitation Theories

Reference in "A short course in Cloud Physics" [16] Precipitation from *stratiform* clouds, which have relatively low liquid water contents, is caused principally by the coalescence ice-crystal process. The clouds last a long time, however, and if cloud persists at altitudes where the temperature is about -15 degree Celsius, the ice-crystal process can lead to precipitation. In an experiment of precipitation development, each level in stratiform cloud has a special role to play in the precipitation process, and the cold upper levels (T = -20 degree Celsius) supplied ice crystal that serve as embryos for precipitation development at lower levels. The cloud at mid-levels (T = -15 degree Celsius) provides the right environment for rapid diffusion growth. Aggregation and accretion proceed most rapidly still lower in the cloud, at temperatures between -10 and 0 degree Celsius. Most of the precipitation growth occurs in these lowest levels.

For *convective* clouds, since liquid water contents are typically higher than in stratiform clouds, therefore less time is available for precipitation growth, and coalescence stands a better chance of producing rain. The lifetime of a convective element (about 20 min.) is also the time needed for precipitation to grow from the observation. In conclusion, the precipitation forming process must begin early in the developing cloud and at a low level. While the precipitation may be initiated by coalescence or the ice-crystal process, depending primarily on the temperature and cloud water content, most precipitation growth is by accretion. Therefore, the mechanisms of precipitation formation are quite different in stratiform and convective clouds. As a useful approximation continuous rain can often be viewed as a steady-state process, in which cloud quantities may vary with height but are constant with time at any given height. Conversely, showers may be approximated as systems in which the cloud properties vary with time but are constant with height at any given time.

3.3.2 Precipitation Processes

There are several types of precipitation processes. The area extent, intensity, and lifetime of a precipitation system are largely controlled by vertical air motions. Thus, it is necessary to classify precipitation as one of two types, depending on the dominant mechanism responsible for the vertical motion:

1. Widespread, *stratiform*, the continuous precipitation related with large scale ascent produced by frontal or topographic lifting or large scale horizontal convergence.

2. Localized, *convective*, the showery precipitation related with cumulus-scale convection in unstable air.

Although the distinction between stratiform and convective precipitation is very complicated, there is a useful classification.

Widespread precipitation invariably shows the fine-scale structure with the most intense precipitation confined to elements with a size of only several kilometers when observed either by radar or rain gauge. But convective precipitation origin can extend over a large area and produce a pattern similar to that of continuous precipitation. Nevertheless, it is usually possible to describe a pattern as either markedly non-uniform (hence convective), which has locally intense regions ranging in size from 1 to 10 km and separation from one another by areas free of precipitation, or rather uniform (hence stratiform). Moreover, the pattern of stratiform precipitation evolves relatively slowly in time and that of convective precipitation changes rapidly. Stratiform rain is produced by nimbostratus clouds, although dissipating cumulus clouds and orographic clouds may generate rain with stratiform structure. Those are what different between convective and stratiform precipitation.