“Artificial Neural Network System for Optical communication System : A sequential Study.”

A Thesis Submitted

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF TECHNOLOGY

IN

Electronics and Communication Engineering-

Wireless Communication and Sensors Network

by

Rajat Gupta

Roll No. 1180454003

Under Supervision of

Asst. Prof. Amit Kumar Singh

BBD University, Lucknow

to the

School of Engineering

BABU BANARASI DAS UNIVERSITY

LUCKNOW

June, 2020

CERTIFICATE

It Is Certified That the Work Contained In This Thesis-1 Entitled "**Artificial Neural Network System for Optical communication System : A sequential Study.**" By **Rajat Gupta** for The Award of Master Of Technology From Babu Banarasi Das University Has Been Carried Out Under My Supervision And That This Work Has Not Been Submitted Elsewhere For A Degree.

Supervisor	Head of Department
Mr. Amit Kumar Singh	Dr. Nitin Jain
(Assistant Professor)	(Associate Professor)
BBD University	BBD University

ABSTRACT

We investigate the risk of overestimating the performance gain when applying neural network based receivers in systems with pseudo random bit sequences or with limited memory depths, resulting in repeated short patterns. We show that with such sequences, a large artificial gain can be obtained which comes from pattern prediction rather than predicting or compensating the studied channel/phenomena. Since using additive white gaussian noise (AWGN) is common to every communication channels, which is the statistically random radio noise characterized by a wide frequency range with regards to a signal in the communications channel. this assignment describes two aspects of telecommunications engineering: i) the basic understanding of MATLAB, and ii) the effect of additive white gaussian noise (AWGN) on the transmitted data using based-band simulation under different values of signal-to-noise (snr) ratio with this tool. in my experiment, I use MATLAB tools to generate the noise sequence, the data sequence, the both data and noise sequence and add them together, recover signal from noisy received data and calculate noise power, calculate SNR and BER and then plot them with different snr values.

ACKNOWLEDGEMENT

It is proud privilege to express a profound sense of gratitude and whole hearted thanks to my respected guide Mr. Amit Kumar Singh (assistant professor), department of electronics & communication engineering, Babu Banarasi Das University, Lucknow, for his expert guidance, invaluable suggestions and encouragement throughout the thesis work.

I owe a great deal of appreciation to Dr. Nitin Jain (HOD), Department of electronics & communication engineering, Babu Banarasi Das University, Lucknow, for his keen interest and valuable guidance during the course of thid work. Their continuous encouragement inspired me throughout the work.

Rajat Gupta

Roll.no: 1180454003

BBD University

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 BACKGROUND: People have used light to transmit information for hundreds of years. However, it was not until the 1960s, with the invention of the laser, that widespread interest in optical (light) systems for data communications began. The invention of the laser prompted researchers to study the potential of fiber optics for data communications, sensing, and other applications. Laser systems could send a much larger amount of data than telephone, microwave, and other electrical systems. The first experiment with the laser involved letting the laser beam transmit freely through the air. Researchers also conducted experiments letting the laser beam transmit through different types of waveguides. Glass fibers, gas-filled pipes, and tubes with focusing lenses are examples of optical waveguides. Glass fibers soon became the preferred medium for fiber optic research.

Initially, the very large losses in the optical fibers prevented coaxial cables from being replaced. Loss is the decrease in the amount of light reaching the end of the fiber. Early fibers had losses around 1,000 dB/km to make them impractical for communications use. In 1969, several scientists concluded that impurities in the fiber material caused the signal loss in optical fibers. The basic fiber material did not prevent the light signal from reaching the end of the fiber. These researchers believed it was possible to reduce the losses in optical fibers by removing the impurities. By removing the impurities, construction of low-loss optical fibers was possible.

In 1970, Corning Glass Works made a multimode fiber with losses under 20 dB/km. This same company, in 1972, made a high silica-core multimode optical fiber with 4dB/km minimum attenuation (loss). Currently, multimode fibers can have losses as low as 0.5 dB/km at wavelengths around 1300 nm.

1.2 INTRODUCTION: Optical fiber has a number of advantages over the copper wire used to make connections electrically. For example, optical fiber, being made of glass or plastic, is immune to electromagnetic interference which is caused by thunderstorms. Also, because light has a much higher frequency than any radio signal we can generate, fiber has a wider bandwidth and can therefore carry more information at one time.

Most telephone company long-distance lines are now of optical fiber. Transmission on optical fiber wire requires repeaters at distance intervals. The glass fiber requiresmore protection within an outer cable than copper. For these reasons and because the installation of any new wiring is labor-intensive, few communities yet have optical fiber wires or cables from the phone company's branch office to local customers.

1.3 FUNDAMENTALS OF FIBERS: The fundamental principle that makes optical fibers possible is total internal reflection. This is described using the ray model of light as shown in following figure.

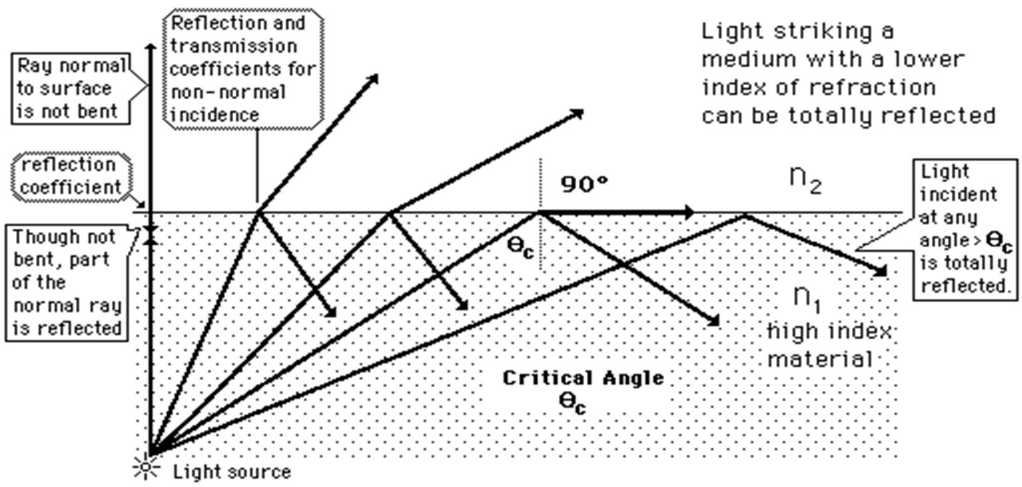


Figure1.1: Total internal Reflection

1.3.1Total Internal Reflection

From Snell's Law we find that refraction (as shown by the dashed line) can only occur when the angle theta1 is large enough. This implies that as the angle is reduced, there must be a point when the light ray is reflected, where theta1 = theta2. The angle where this happens is known as the critical angle and is:

$$\sin \theta = \frac{n2}{n1}$$

1.3.2 Transmitters

Fiber optic transmitters are devices that include an LED or laser source, and signal conditioning electronics, to inject a signal into fiber. The modulated light may be turned on or off, or may be linearly varied in intensity between two predetermined levels. Light

Emitting Diodes (LEDs) have relatively large emitting areas and as a result are not as good light sources as laser diodes. However, they are widely used for short to moderate transmission distances because they are much more economical. Laser diodes can couple many times more power to optical fiber than LEDs. They are primarily used for applications that require the transmission of signals over long distances.

Important performance specifications to consider when searching for fiber optic transmitters include data rate, transmitter rise time, wavelength, spectral width, and maximum optical output power. Data rate is the number of data bits transmitted in bits per second. Data rate is a way of expressing the speed of the transceiver. In the approximation of a step function, the transmitter rise time is the time required for a signal to change from a specified 10% to 90% of full power. Rise time is a way of expressing the speed of the transmitter. Wavelength refers to the output wavelength of the transceiver. The spectral width refers to the spectral width of the output signal.

1.3.3 Receivers

Fiber optic receivers are instruments that convert light into electrical signals. They contain a photodiode semiconductor, signal conditioning circuitry, and an amplifier. Fiber optic receivers use three types of photodiodes: positive-negative (PN) junctions, positive-intrinsic-negative (PIN) photodiodes, and avalanche photodiodes (APD). PIN photodiodes have a large, neutrally-doped region between the p-doped and n-doped regions. APDs are PIN photodiodes that operate with high reverse-bias voltages. In short wavelength fiber optic receivers (400 nm to 1100 nm), the photodiode is made of silicon (Si). In long wavelength systems (900 nm to 1700 nm), the photodiode is made of indium gallium arsenide (InGaAs). With low-impedance amplifiers, bandwidth and receiver noise decrease with resistance. With trans-impedance amplifiers, the bandwidth of the receiver is affected by the gain of the amplifier. Typically, fiber optic receivers include a removable adaptor for connections to other devices. Data outputs include transistor-transistor logic (TTL), emitter-coupled logic (ECL), video, radio frequency (RF), and complementary metal oxide semiconductor (CMOS) signals. Also, it uses many types of connectors.

1.3.4 Fiber

Fiber is the medium to guide the light form the transmitter to the receiver. It is classified into two types depending on the way the light is transmitted: multimode fiber and single-mode fiber. Fiber can be classified in two ways:

1.4 BASED ON THE NUMBER OF MODES

1.4.1 Single Mode fiber

when a fiber wave-guide can support only the HE11 mode, it is referred to as a single mode wave-guide. In a step index structure this occurs when the wave-guide is operating at v<2.4 where v is dimensionless number which relates the propagating in the cladding. These single mode fibers have small size and low dopant level (typically 0.3% to 0.4% index elevation over the lading index.)

In high silica fibers the wave-guide and the material dispersion are often of opposite signs. This fact can be used conveniently to achieve a single mode fiber of extremely large bandwidth.

Reduced dopant level results in lower attenuation than in multimode fibers. A single mode wave guide with its large and fully definable bandwidth characteristics is an obvious candidate for long distance, high capacity transmission applications.

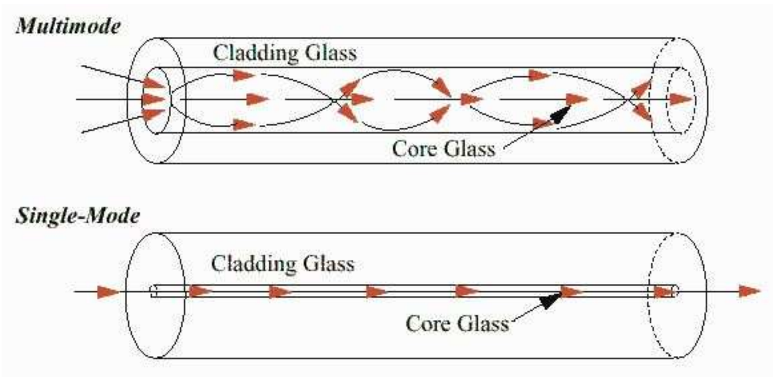


Figure 1.2: Single and Multimode fiber

1.4.2 Multimode fiber

It is a fiber in which more than one mode is propagating at the system operating wavelength. Multimode fiber system does not have the information carrying capacity of single mode fibers.

However they offer several advantages for specific system. The larger core diameters result in easier splicing of fibers. Given the larger cores, higher numerical apertures, and typically shorter link distances, multimode systems can less expensive light sources such as LEDs.

Multimode fibers have numerical apertures that typically range from 0.2 to 0.29 and have core size that range from 35 to 100 micro-meters.

1.5 BASED ON REFRACTIVE INDEX

1.5.1 Step index fiber

The step index (SI) fiber consists of a central core whose refractive index is n1, surrounded by a lading whose refractive index is n2, lower than that of core. Because of an abrupt index change at the core cladding interface such fibers are called step index fibers.

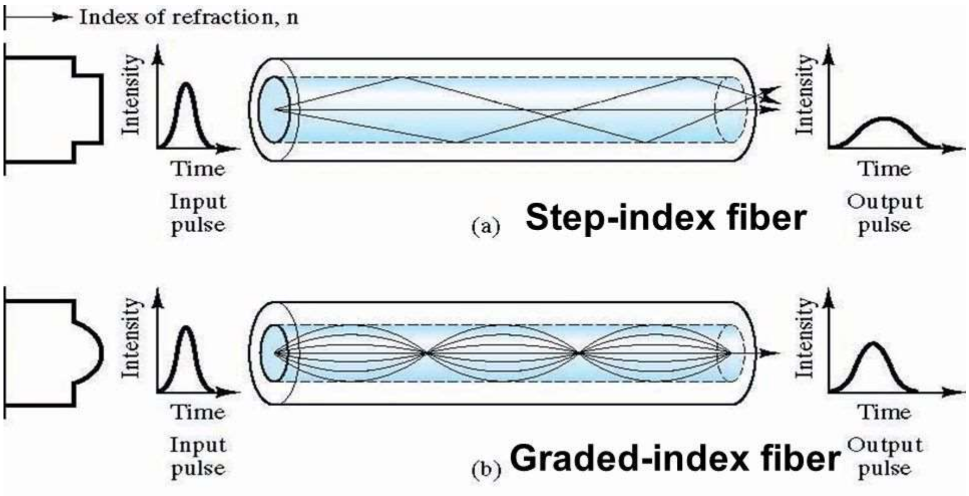


Figure 1.3: Step index and Graded Index Fiber

1.5.2 Graded index fibers

The refractive index of the core in graded index fiber is not constant, but decreases gradually from its maximum value n1 to minimum value n2 at the core-cladding interface. The ray velocity changes along the path because of variations in the refractive index. The ray propagating along the fiber axis takes the shortest path but travels most slowly, as the index is largest along this path in medium of lower refractive index where they travel faster. It is therefore possible for all rays to arrive together at the fiber output by a suitable choice of refractive index profile.

1.6 ADVANTAGES OVER CONVENTIONAL CABLES

- Wide Bandwidth: Optical fibers offer greater bandwidth due to the use of light as carrier. The frequency range used for glass fiber communication extends from 2*e14Hz to 4*e14Hz. Hence optical fibers are suitable for high speed, large capacity telecommunication lines.

- Low Loss: In a coaxial cable attenuation increases with frequency. The higher the frequency of information signals the greater the loss, whereas in an optical fiber the attenuation is independent of frequency. They offer a loss of 0.2 dBm/km, allowing repeater separation upto 50Km or more.

- Freedom from electromagnetic interference: Optical fibers are not affected by interference originating from power cables, railways and radio waves. They do not limit unwanted radiation and no cross talk between fibers exists. These fibers make an ideal transmission medium when EMI (Electro Magnetic Immunity) is increased.

- Non conductivity: Optical fibers are non-conductive and are not effective by strong electromagnetic interference such as lighting. These are usable in explosive environment.

- Small diameters and less weight: Even multi fiber optical cables have a small diameter and are light weight, and flexible optical fiber cables permit effective utilization of speech and can also be applicable to long distance use are easier to handle and install than conventional cables.

- Security: Fiber optic is a highly source transmission medium. It does not radiate energy that can be received by a nearby antenna, and it is extremely difficult to tap a fiber and virtually impossible to make the tap undetected.

- Safety: Fiber is a dielectric and does not carry electricity. It presents no sparks or fire hazards. It does not cause explosions, which occur due to faulty copper cable.

1.7 DISADVANTAGES OF OPTICAL FIBER

- **Cost:** Cables are expensive to install but last longer than copper cable.
- **Transmission:** Transmission on optical fiber requires repeater at distance intervals.

- **Fragile:** fibers can broken and have transmission loses and wrapped around curves only a few centimeters radius.
- **Protection:** optical fiber requires more protection around the cable compared to copper.

1.8 APPLICATIONS

Due to the advantages of fiber optic over the traditional connectivity networks, networks are being changed to the new technology of fiber optic. Here is some applications use fiber optics for the communication:

- Long Haul telecommunication systems on land and at sea to carry many simultaneous telephone calls (or other signals) over long distances. These include ocean spanning submarine cables and national backbone networks for telephone and computer data transmission.
- Interoffice trunks that carry many telephone conversations simultaneously between local and regional switching facilities.
- Connections between the telephone N/W and antennas for mobile telephone service.
- Links among computers and high resolution video-terminals used for such purposes as computer aided design.
- Transmission of signals within ships and aircraft.
- Local area Networks operating at high speeds or over large areas, and backbone systems connecting slower local area Networks.
- High speed interconnections between computer and peripherals devices, or between computers, or even within segments of single large.

1.9 BLOCK SCHEMATIC OF COMMUNICATION SYSTEM

An optical fiber communication system is similar in basic concept to any type of communication system. A block schematic of a general communication system is shown in Figure 1.1, the function of which is to convey the signal from the information source over the transmission medium to the destination. The communication system therefore consists of a transmitter or modulator linked to the information source, the transmission medium, and a receiver or demodulator at the destination point. In electrical communications the information source provides an electrical signal, usually derived from a message signal which is not electrical (e.g. sound), to a transmitter comprising electrical and electronic components which converts the signal into a suitable form for propagation over the transmission medium. This is often achieved by modulating a carrier, which, as mentioned previously, may be an electromagnetic wave. The transmission medium can consist of a pair of wires, a coaxial cable or a radio link through free space down which the signal is transmitted to the receiver, where it is transformed into the original electrical information signal (demodulated) before being passed to the destination. However, it must be noted that in any transmission medium the signal is attenuated, or suffers loss, and is subject to degradations due to contamination by random signals and noise, as well as possible distortions imposed by mechanisms within the medium itself. Therefore, in any communication system there is a maximum permitted distance between the transmitter and the receiver beyond which the system effectively ceases to give intelligible communication.

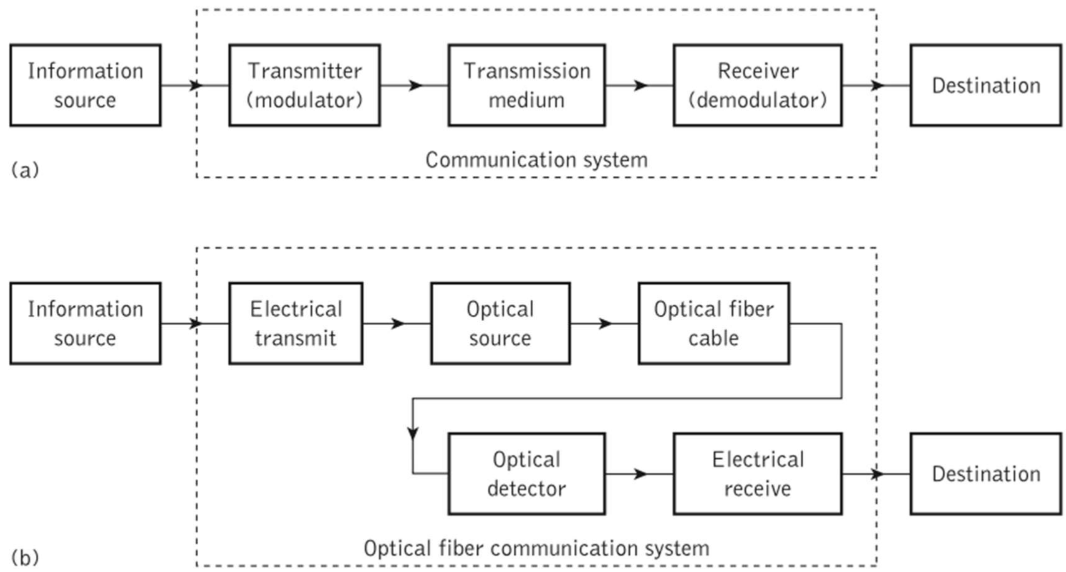


Figure 2.1: Block schematic of communication system

Since the development of low-loss optical fiber cables in the 1970s, optical communications became one of the most popular methods of communication.

1.10 TYPICAL DIGITAL OPTICAL FIBER LINK:

The optical carrier may be modulated using either an analog or digital information signal. In the system shown in Figure 1.2 analog modulation involves the variation of the light emitted from the optical source in a continuous manner. With digital modulation, however, discrete changes in the light intensity are obtained (i.e. on–off pulses). Although often simpler to implement, analog modulation with an optical fiber communication sys- tem is less efficient, requiring a far higher signal-to-noise ratio at the receiver than digital modulation. Also, the linearity needed for analog modulation is not always provided by semiconductor optical sources, especially at high modulation frequencies. For these reasons, analog optical fiber communication links are generally limited to shorter distances and lower bandwidth operation than digital links.

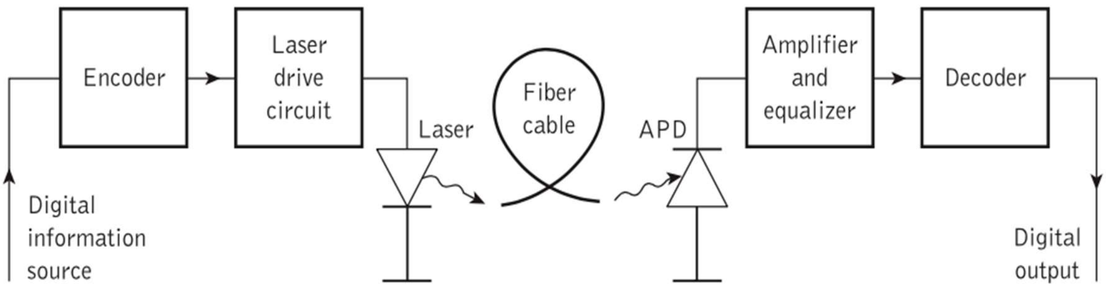


Figure 2.2: Typical digital Optical Fiber link

Figure 2.2 shows a block schematic of a typical digital optical fiber link. Initially, the input digital signal from the information source is suitably encoded for optical transmission. The laser drive circuit directly modulates the intensity of the semiconductor laser with the encoded digital signal. Hence a digital optical signal is launched into the optical fiber cable. The avalanche photodiode (APD) detector is followed by a front-end amplifier and equalizer or filter to provide gain as well as linear signal processing and noise band- width reduction. Finally, the signal obtained is decoded to give the original digital information. The various elements of this and alternative optical fiber system configurations are discussed in detail in the following chapters. However, at this stage it is instructive to consider the advantages provided by light wave communication via optical fibers in comparison with other forms of line and radio communication which have brought about the extensive use of such systems in many areas throughout the world.

1.11 A MORE COMPLECATED NEURON

The previous neuron doesn't do anything that conventional computers don't do already. A more sophisticated neuron (figure 2.7) is the McCulloch and Pitts model (MCP). The difference from the previous model is that the inputs are 'weighted', the effect that each input has at decision making is dependent on the weight of the particular input. The weight of an input is a number which when multiplied with the input gives the weighted input. These weighted inputs are then added together and if they exceed a pre-set threshold value, the neuron fires. In any other case the neuron does not fire.

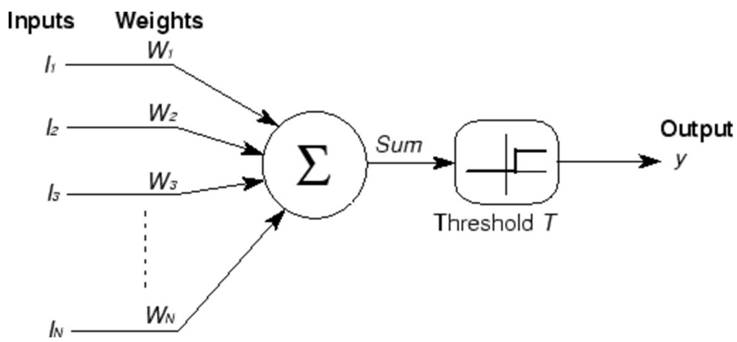


Figure 2.3: The McCulloch-Pitts Model

In mathematical terms, the neuron fires if and only if;

$$\mathbf{X1W1 + X2W2 + X3W3 + ... > T}$$

The addition of input weights and of the threshold makes this neuron a very flexible and powerful one. The MCP neuron has the ability to adapt to a particular situation by changing its weights and/or threshold. Various algorithms exist that cause the neuron to 'adapt'; the most used ones are the Delta rule and the back error propagation. The former is used in feed-forward networks and the latter in feedback networks.

1.11.1 Feed-forward networks

Feed-forward ANNs (figure 2.8) allow signals to travel one way only; from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. This type of organisation is also referred to as bottom-up or top-down.

1.11.2 Feedback networks

Feedback networks (figure 2.8) can have signals travelling in both directions by introducing loops in the network. Feedback networks are very powerful and can get extremely complicated. Feedback networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organisation.

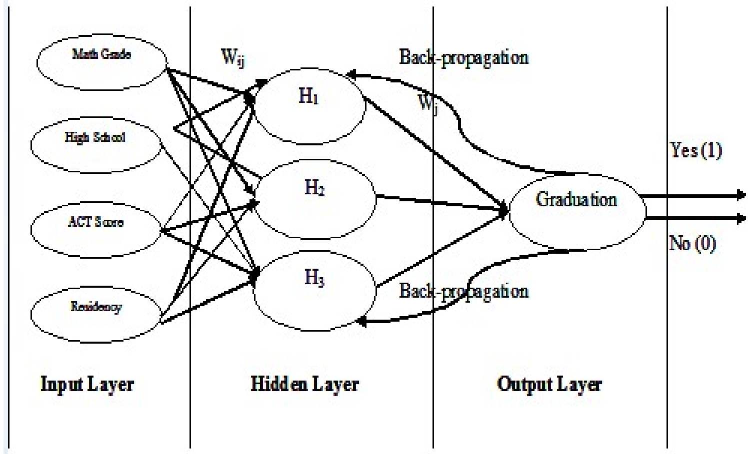


Figure 2.4: An example of simple feedforward network.

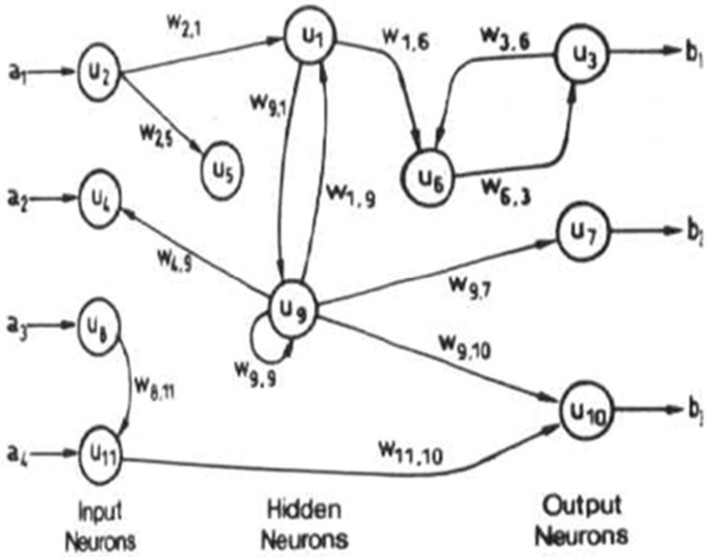


Figure 2.5: An example of complicated neural network.

1.12 ACTIVATION FUNCTION

Activation Functions are basically the transfer function which is output from a artificial neuron and it send signals to the other artificial neuron. There are four form of Activation Functions Threshold, Piecewise-Linear, Sigmoid and Gaussian all are different from each other. In Below figures you can see the Activation function with its demonstration.

1.12.1 Threshold:

$$f(x)=\begin{cases}0 \text{ if } 0>x \\ 1 \text{ if } x\geq 0\end{cases}$$

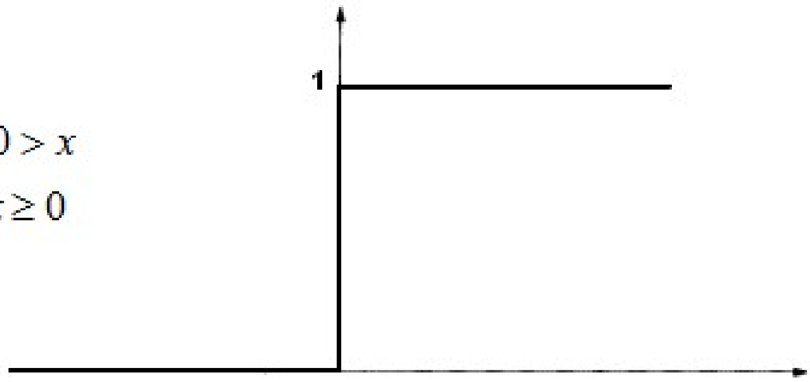


Figure 2.6: Threshold

1.12.2 Signum Function:

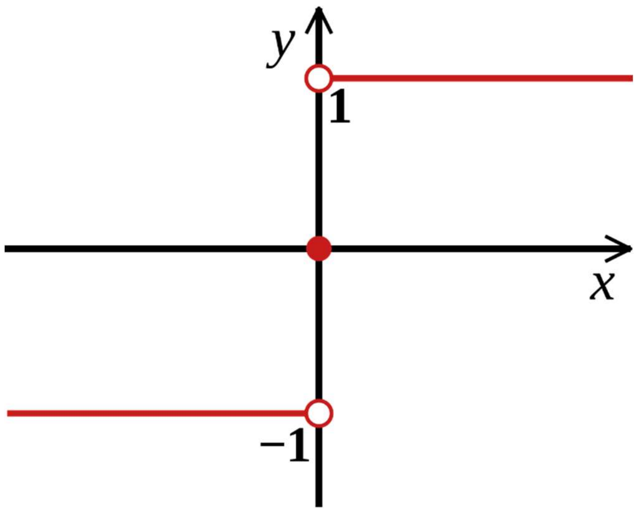


Figure 2.7: Signum Function

1.12.3 Piece-wise Linear Function

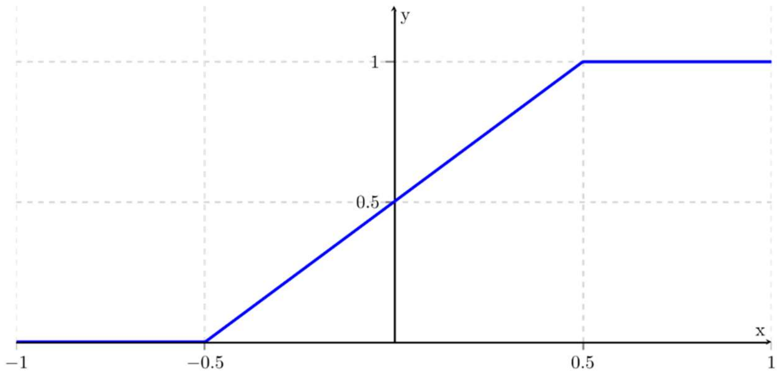


Figure 2.8: Piece-wise Linear Function

1.12.4 Signoidal Function

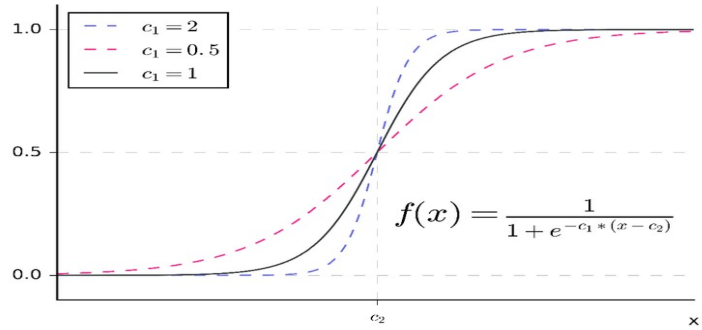


Figure 2.9: Signoidal Function

1.13 NETWORK LAYERS

The commonest type of artificial neural network consists of three groups, or layers, of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. (see Figure 2.8)

- The activity of the input units represents the raw information that is fed into the network.
- The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units.
- The behaviour of the output units depends on the activity of the hidden units and the weights between the hidden and output units.

This simple type of network is interesting because the hidden units are free to construct their own representations of the input. The weights between the input and hidden units determine when each hidden unit is active, and so by modifying these weights, a hidden unit can choose what it represents.

We also distinguish single-layer and multi-layer architectures. The single-layer organisation, in which all units are connected to one another, constitutes the most general case and is of more potential computational power than hierarchically structured multi-layer organisations. In multilayer networks, units are often numbered by layer, instead of following a global numbering.

1.14 PERCEPTRONS

The most influential work on neural nets in the 60's went under the heading of 'perceptrons' a term coined by Frank Rosenblatt. The perceptron (figure 4.4) turns out to be an MCP model (neuron with weighted inputs) with some additional, fixed, pre--processing. Units labelled A1, A2, Aj, Ap are called association units and their task is to extract specific, localized featured from the input images. Perceptrons mimic the basic idea behind the mammalian visual system. They were mainly used in pattern recognition even though their capabilities extended a lot more.

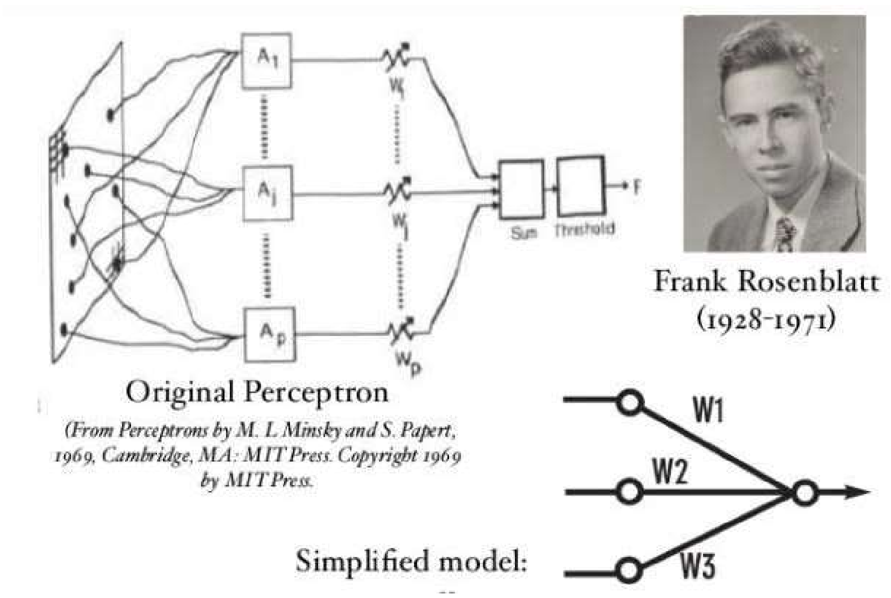


Figure 2.10: Perceptrons

In 1969 Minsky and Papert wrote a book in which they described the limitations of single layer Perceptrons.

The impact that the book had was tremendous and caused a lot of neural network researchers to loose their interest. The book was very well written and showed mathematically that single layer perceptrons could not do some basic pattern recognition operations like determining the parity of a shape or determining whether a shape is connected or not. What they did not realized, until the 80's, is that given the appropriate training, multilevel perceptrons can do these operations.

1.15 THE LEARNING PROCESS

The memorization of patterns and the subsequent response of the network can be categorised into two general paradigms:

- **associative mapping** in which the network learns to produce a particular pattern on the set of input units whenever another particular pattern is applied on the set of input units. The associtive mapping can generally be broken down into two mechanisms:

- **auto-association**: an input pattern is associated with itself and the states of input and output units coincide. This is used to provide pattern completition, ie to produce a pattern whenever a portion of it or a distorted pattern is presented. In the second case, the network actually stores pairs of patterns building an association between two sets of patterns.

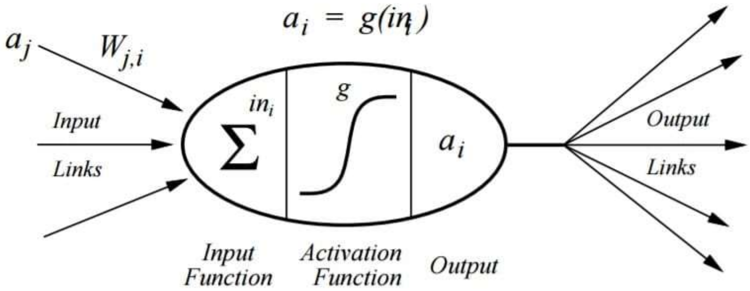
- **hetero-association**: is related to two recall mechanisms:

  - **nearest-neighbour recall**, where the output pattern produced corresponds to the input pattern stored, which is closest to the pattern presented, and

  - **interpolative recall**, where the output pattern is a similarity dependent interpolation of the patterns stored corresponding to the pattern presented. Yet another paradigm, which is a variant associative mapping is classification, i.e. when there is a fixed set of categories into which the input patterns are to be classified.

- **regularity detection** in which units learn to respond to particular properties of the input patterns. Whereas in asssociative mapping the network stores the relationships among patterns, in regularity detection the response of each unit has a particular 'meaning'. This type of learning mechanism is essential for feature discovery and knowledge representation.

Every neural network posseses knowledge which is contained in the values of the connections weights. Modifying the knowledge stored in the network as a function of experience implies a learning rule for changing the values of the weights.



$$a_i = g(\sum_j W_{j,i}a_j)$$

Figure 2.11: Learning Process

Information is stored in the weight matrix W of a neural network. Learning is the determination of the weights. Following the way learning is performed, we can distinguish two major categories of neural networks:

- fixed networks in which the weights cannot be changed, i.e. dW/dt=0. In such networks, the weights are fixed a priori according to the problem to solve.

- adaptive networks which are able to change their weights, i.e. dW/dt not= 0.

All learning methods used for adaptive neural networks can be classified into two major categories:

- Supervised learning which incorporates an external teacher, so that each output unit is told what its desired response to input signals ought to be.

- During the learning process global information may be required. Paradigms of supervised learning include error-correction learning, reinforcement learning and stochastic learning. An important issue conserning supervised learning is the problem of error convergence, ie the minimisation of error between the desired and computed unit values. The aim is to determine a set of weights which minimizes the error. One well-known method, which is common to many learning paradigms is the least mean square (LMS) convergence.

- Unsupervised learning uses no external teacher and is based upon only local information. It is also referred to as self-organisation, in the sense that it self-organize data presented to the network and detects their emergent collective properties. Paradigms of unsupervised learning are Hebbian lerning and competitive learning.

- From Human Neurons to Artificial Neurons the aspect of learning concerns the distinction or not of a seperate phase, during which the network is trained, and a subsequent operation phase. We say that a neural network learns off-line if the learning phase and the operation phase are distinct. A neural network learns on-line if it learns and operates at the same time. Usually, supervised learning is performed off-line, whereas unsupervised learning is performed on-line.

An Example to illustrate the above teaching procedure:

Assume that we want a network to recognize hand-written digits. We might use an array of, say, 256 sensors, each recording the presence or absence of ink in a small area of a single digit. The network would therefore need 256 input units (one for each sensor), 10 output units (one for each kind of digit) and a number of hidden units.

For each kind of digit recorded by the sensors, the network should produce high activity in the appropriate output unit and low activity in the other output units.

To train the network, we present an image of a digit and compare the actual activity of the 10 output units with the desired activity.

We then calculate the error, which is defined as the square of the difference between the actual and the desired activities. Next we change the weight of each connection so as to reduce the error.

We repeat this training process for many different images of each different images of each kind of digit until the network classifies every image correctly.

To implement this procedure we need to calculate the error derivative for the weight (EW) in order to change the weight by an amount that is proportional to the rate at which the error changes as the weight is changed. One way to calculate the EW is to perturb a weight slightly and observe how the error changes. But that method is inefficient because it requires a separate perturbation for each of the many weights.

Another way to calculate the EW is to use the Back-propagation algorithm which is described below, and has become nowadays one of the most important tools for training neural networks. It was developed independently by two teams, one (Fogelman-Soulie, Gallinari and Le Cun) in France, the other (Rumelhart, Hinton and Williams) in U.S.

The Back-Propagation Algorithm

In order to train a neural network to perform some task, we must adjust the weights of each unit in such a way that the error between the desired output and the actual output is reduced.

This process requires that the neural network compute the error derivative of the weights (EW). In other words, it must calculate how the error changes as each weight is increased or decreased slightly. The back propagation algorithm is the most widely used method for determining the EW.

The back-propagation algorithm is easiest to understand if all the units in the network are linear. The algorithm computes each EW by first computing the EA, the rate at which the error changes as the activity level of a unit is changed.

For output units, the EA is simply the difference between the actual and the desired output.

To compute the EA for a hidden unit in the layer just before the output layer, we first identify all the weights between that hidden unit and the output units to which it is connected. We then multiply those weights by the EAs of those output units and add the products. This sum equals the EA for the chosen hidden unit. After calculating all the EAs in the hidden layer just before the output layer, we can compute in like fashion the EAs for other layers, moving from layer to layer in a direction opposite to the way activities propagate through the network.

This is what gives back propagation its name. Once the EA has been computed for a unit, it is straight forward to compute the EW for each incoming connection of the unit. The EW is the product of the EA and the activity through the incoming connection.

Note that for non-linear units, the back-propagation algorithm includes an extra step. Before back-propagating, the EA must be converted into the EI, the rate at which the error changes as the total input received by a unit is changed.

1.16 RECENT ADVANCES AND FUTURE APPLICATIONS OF NNs INCLUDE:

Integration of fuzzy logic into neural networks

Fuzzy logic is a type of logic that recognizes more than simple true and false values, hence better simulating the real world. For example, the statement today is sunny might be 100% true if there are no clouds, 80% true if there are a few clouds, 50% true if it’s hazy, and 0% true if rains all day. Hence, it takes into account concepts like -usually, somewhat, and sometimes. Fuzzy logic and neural networks have been integrated for uses as diverse as automotive engineering, applicant screening for jobs, the control of a crane, and the monitoring of glaucoma.

Pulsed neural networks

Most practical applications of artificial neural networks are based on a computational model involving the propagation of continuous variables from one processing unit to the next.

In recent years, data from neurobiological experiments have made it increasingly clear that biological neural networks, which communicate through pulses, use the timing of the pulses to transmit information and perform computation. This realization has stimulated significant research on pulsed neural networks, including theoretical analyses and model development, neurobiological modeling, and hardware implementation."

Hardware specialized for neural networks

Some networks have been hardcoded into chips or analog devices? this technology will become more useful as the networks we use become more complex. The primary benefit of directly encoding neural networks onto chips or specialized analog devices is SPEED! NN hardware currently runs in a few niche areas, such as those areas where very high performance is required (e.g. high energy physics) and in embedded applications of simple, hardwired networks (e.g. voice recognition).

Many NNs today use less than 100 neurons and only need occasional training. In these situations, software simulation is usually found sufficient When NN algorithms develop to the point where useful things can be done with 1000's of neurons and 10000's of synapses, high performance NN hardware will become essential for practical operation.

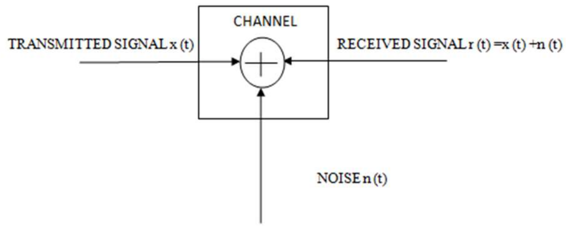
Improvement of existing technologies

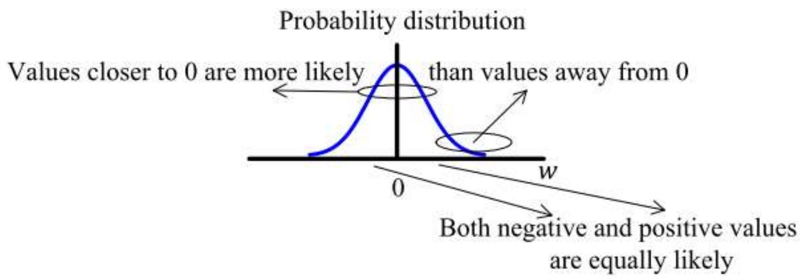
All current NN technologies will most likely be vastly improved upon in the future. Everything from handwriting and speech recognition to stock market prediction will become more sophisticated as researchers develop better training methods and network architectures.

NNs might, in the future, allow:

robots that can see, feel, and predict the world around them improved stock prediction common usage of self-driving cars composition of music handwritten documents to be automatically transformed into formatted word processing documents trends found in the human genome to aid in the understanding of the data compiled by the Human Genome Project self-diagnosis of medical problems using neural networks.

As AWGN channel is used due to uniform power spectral density over the entire frequency range of interest. And the PDF of it is same as gaussian surface.





However, to limit our scope, in this book we're going to concentrate on the more widely used feedforward networks. Author in [3] showed that a large deep LSTM with a limited vocabulary can outperform a standard STM-based system whose vocabulary is unlimited on a large scale MT task. The success of our simple LSTM-based on MT suggest that it should do well on many other sequence learning problem, provided they have enough training data.

In this paper we present a general end-to-end approach to sequence learning to make minimal assumptions on the sequence structure. In [4] author stated that We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes.

On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art.

The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax.

To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation.

To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called “dropout” that proved to be very effective. In [5] An ideal face classifier would recognize faces in accuracy that is only matched by humans.

The underlying face descriptor would need to be invariant to pose, illumination, expression, and image quality. It should also be general, in the sense that it could be applied to various populations with little modifications, if any at all. In addition, short descriptors are preferable, and if possible, sparse features. Certainly, rapid computation time is also a concern. We believe that this work, which departs from the recent trend of using more features and employing a more powerful metric learning technique, has addressed this challenge, closing the vast majority of this performance gap.

Our work demonstrates that coupling a 3D model-based alignment with large capacity feedforward models can effectively learn from many examples to overcome the drawbacks and limitations of previous methods.

In author [6] we stated that it gives a brief description of deep learning for the nonlinear classifier and the background algorithm with an example of movie recommendations as Gravity and ask for rating and all are giving different points than they have to calculate the data using stochastic gradient descent to minimize a decision function.

The main problem encounters that ANNs tends to overfit, especially when the training set is small. To overcome overfitting some methods are use unlabeled data to train, penalize the weights by adding a penalty, use dropout.

In another paper [8] we have discussed various NN architectures and learning methods for solving the problem of channel equalization. The main drawback of the NN equalizers is the large computational complexity due to extensive training. The MLP network is simple to implement, but usually requires long training time.

The main limitation of FLANN structure is that as the number of nodes in the input signal space is increased, the computational complexity increases. RBF-based NN equalizers are an attractive alternative and have successfully being applied for blind equalization. RNN-based equalizers, generalized as IIR filters, outperform feedforward NNs, including MLP, RBF, and FLANN. They are especially suitable for equalization of fading channels.

In [9] author described the investigation of application of ANN to adaptive and blind equalization problems (i) to introduce a new realization structure of a multilayered preceptrons with back propagation training algorithm, (ii) to demonstrate the performance of self organizing maps as blind equalizers. Our work provides the first application evidence that MLP is not only a specified method for pattern classification but also potentially capable for a broader class of problem.

In [10] they extend the capabilities of neural networks by coupling them to external memory resources, which they can interact with by attentional processes.

The combined system is analogous to a Turing Machine or Von Neumann architecture but is differentiable end-to- end, allowing it to be efficiently trained with gradient descent. Preliminary results demonstrate that Neural Turing Machines can infer simple algorithms such as copying, sorting, and associative recall from input and output examples. In [11] The gain of the compensation method are therefore investigated by varying the parameters of the AC function describing XPM-induced polarization scattering and phase noise.

It is shown that an increase in the nonlinear tolerance of more than 2 dB is achievable for 32 Gbaud QPSK and 16- quadratic-amplitude modulation (QAM). It is also reviewed how laser rate equations can be formulated within the nonlinear state- space framework which allows for tracking of non-Lorentzian laser phase noise line-shapes.

It is experimentally demonstrated for 28 Gbaud 16-QAM signals that if the laser phase noise shape strongly deviates from the Lorentzian, phase noise tracking algorithms employing rate equation-based SSM result in a significant performance improvement (>8 dB) compared to traditional approaches using digital phase-locked loop.

Finally, Gaussian mixture model is reviewed and employed for nonlinear phase noise compensation and characterization of nano scale devices structure variations. In another [12-14] A novel low-complexity ANN-based NLE has been proposed for CO-OFDM systems. ANN-NLE proved to be a robust nonlinearity DSP technique for up to 80-Gb/s 16-QAM CO-OFDM systems. For 80-Gb/s transmission over 1000-km uncompensated link, ANN-NLE outperforms in terms of Q-factor, LE and IVSTF-NLE by 3 dB and 1 dB, respectively.

This letter should trigger the implementation of nonlinear ANN-based equalizers in next generation core networks. In [15-16] In this paper, a novel RBFNN-NLE for 16-QAM CO- OFDM system has been proposed. We show that our proposed algorithm can reduce NP by 4.5 dB at 80-Gbps with only 3% training overhead.

The results show that at an optimum P_{in} of -3 dBm, we achieve a transmission reach of up to 1200 km for Q-factor of 8.7 dB with our proposed RBFNN-NLE. Hence, the operating data rate and reach are both increased compared to ANN-NLE.

And in [17] A new approach for EPD with a DML based on a neural net- work was introduced. With that approach, a maximum trans- mission distance of 350-km SSMF was achieved by simulations with standard 10-Gb/s DML parameters. Experimentally, a transmission distance of 190-km SSMF was obtained using a commercially available 2.5-Gb/s DML.

In [18] experimentally demonstrated a 4 km 32 GBd PAM-8 IM/DD link with digital equalization based on neural network. We show that the NN-based equalizer is able to accurately reconstruct the received signal suffering from strong ISI. Enhanced equalization performance of the NNE is observed comparing to the conventional linear FFE with as many as 21 taps.

This study indicates that NNE has the potential to be applied for system implementation in optical interconnects and is worthy of further detailed explorations. In [20-21] We investigated equalization performance of a three-layer NN to compensate nonlinear distortion in optical communication systems. Our numerical simulation of 16QAM transmission showed that the NN can efficiently compensate the nonlinear distortion caused by SPM, and improves the performance of BER and EVM. In [22] they propose a simple and cost-effective technique for modulation format identification (MFI) in next-generation heterogeneous fiber-optic networks using an artificial neural network (ANN) trained with the features extracted from the asynchronous amplitude histograms (AAHs).

Results of numerical simulations conducted for six different widely-used modulation formats at various data rates demonstrate that the proposed technique can effectively classify all these modulation formats with an overall estimation accuracy of 99.6% and also in the presence of various link impairments. The proposed technique employs extremely simple hardware and digital signal processing (DSP) to enable MFI and can also be applied for the identification of other modulation formats at different data rates without necessitating hardware changes.

In [23] Parametric Asynchronous Eye Diagram (PAED) is a straightforward and cost-effective alternative to synchronous eye diagram, when the purpose is to visually evaluate the quality of the signal. PAED, together with artificial neural networks showed to be a valid solution in the mon- itoring of different types of traffic with different modulation formats and bit rates, including 10 Gbit/s NRZ, 40 Gbit/s NRZ-QPSK and mixed traffic with different OOK modulation formats and bit rates traveling through the network, such as 10 Gbit/s NRZ, 10 Gbit/s RZ, 20 Gbit/s RZ and 20 Gbit/s NRZ.

CHAPTER 2
LITERATURE REVIEW

Literature Survey is an essential task to review the significant points of the present knowledge together with substantive findings as well as Theoretical and Methodological offerings to Optical fiber network system with Artificial Neural Network. In order to become familiar with current literature on Optical fiber network system with Artificial Neural Network**, i**nformation is assembled from Books, Journals, Conference Proceedings etc.

Danshi Wang et.al.[1] proposed a new combined approach that is ANN with Coherent optical fiber system. ANN technique decreases the noise that comes during the transmission time. In optical fiber transmission there are various types of interrupts arises. To resolve these interrupts ANN, produce effective result.

.

Elias Giacoumidis et.al.[2] presented a novel NLE based on Artificial Neural Network.
Q-factor is used to perform the following implementation. ANN neural is used to sweep away the NLE (Non-Linear Equalizer) performance. ANN improves the optical fiber system performance.

Alex Krizhevsky et.al.[3] proposed a deep convolutional neural network concepts that help to achieve effective performance. Used large dataset that verify the proposed system performance. More than one number of convolutional layers are used to perform proposed operation.

Marcia Peng et.al.[4] discussed the neural networks applications. First is used to present a multilayer perception with backpropagation training and in second self -organizing maps as blind equalizers.

Simone Gaiarin et.al.[5] presented a high-speed optical interconnection network with neural network. Using this system performance of data transmission is increased. For performance comparison Enhanced equalization is used.

Hua Wang et.al.[6] proposed a new integrated approach that QKD (Quantum Key Distribution) with Optical Fiber. Two primary goals are targeted to achieve using it is cost-effectiveness and scalability. Improve the transmission efficiency using this approach.

Kevin Cushon et.al. [7] present a low-density parity check (LDPC) for optical networks. Application specific integrated circuit synthesis is also discussed. Provides better efficiency compare to other existing techniques.

Alex Alvarado et.al.[8] present a Forward Error Correction (FEC) that is used to design Optical Fiber Network. It is used to reduce Bit Error Rate (BER). Extensive optical fields are used to carry out in both cases i.e. linear and non-linear transmission.

Stefan Warm et.al.[9] presents a Directly Modulated Laser (DML). This approach is based on Artificial Neural Network (ANN). Experiment bon single mode fiber. Improve the efficiency of current approach.

F.N. Khan et.al.[10] proposed a cost-effective technique for Modulation Format Identification (MFI) with Artificial Neural Network (ANN). The system trained by using Asynchronous Amplitude Histogram (AAH). Provides effective result in optical network data transmission.

TABLE 1: INFERENCE DRAWN OUT OF LITERATURE REVIEW

S. No	Authors Name	Publication and Years	Methodology Used	Inference Drawn
1	Alex Alvarado et.al	**IEEE ,2015**	Used Forward Error Correction (FEC) and Bit Error Correction (BER).	• FEC provides an effective approach required in optical transmission. • A better predictor for BER. • Sort out the previous issues using FEC.
2.	Hua Wang et.al.	**IEEE,2019**	Used novel technique Quantum Key Distribution (QKD).	• Secret key flow model (SKFM) is used. • Present Secret Key Effective Strategy (SKES) •
3.	Simone Gaiarin et.al.	**ACP,2016**	Used optical interconnection link with Neural	• Present Intensity modulation and Direct Detection system.

			Network Equalization.	• Pulse Amplitude Modulation (PAM) is used. • It helps to find out the effective result.
4	Stefan Warm et.al.	**IEEE,2009**	Use Directly Modulated Laser (DML)Technique.	• Apply Electronic Dispersion Precompensation to DML. • Use new approach that is based on ANN. • Use single mode fiber for data transmission.
5.	Elias Giacoumidis et.al.	**OSA,2015**	Use Q-factor in terms of optic fiber.	• Use Non-Linear Equalizer with ANN. • NLE improves performance efficiency. • Performance of the system based on many factors, in this paper try to focus on these.

CHAPTER 3

PROBLEM FORMULATION AND SOLUTION METHODOLOGY

3.1 OBJECTIVE

Enhancement of performance gain using Pseudo-random bit sequence in NNs based Optical communication Receiver.

3.2 PROBLEM FORMULATION

When pseudo random bit sequence added with AWGN noise then the obtained bit error rate is very high on comparing with hard limiter or threshold but when neural network is applied in the system then obtained BER is less and SNR is high so we can enhance the performance gain of the system using Neural Network.

3.3 SOLUTION METHODOLOGY

We investigate the risk of overestimating the performance gain when applying neural network-based receivers in systems with pseudo random bit sequences or with limited memory depths, resulting in repeated short patterns.

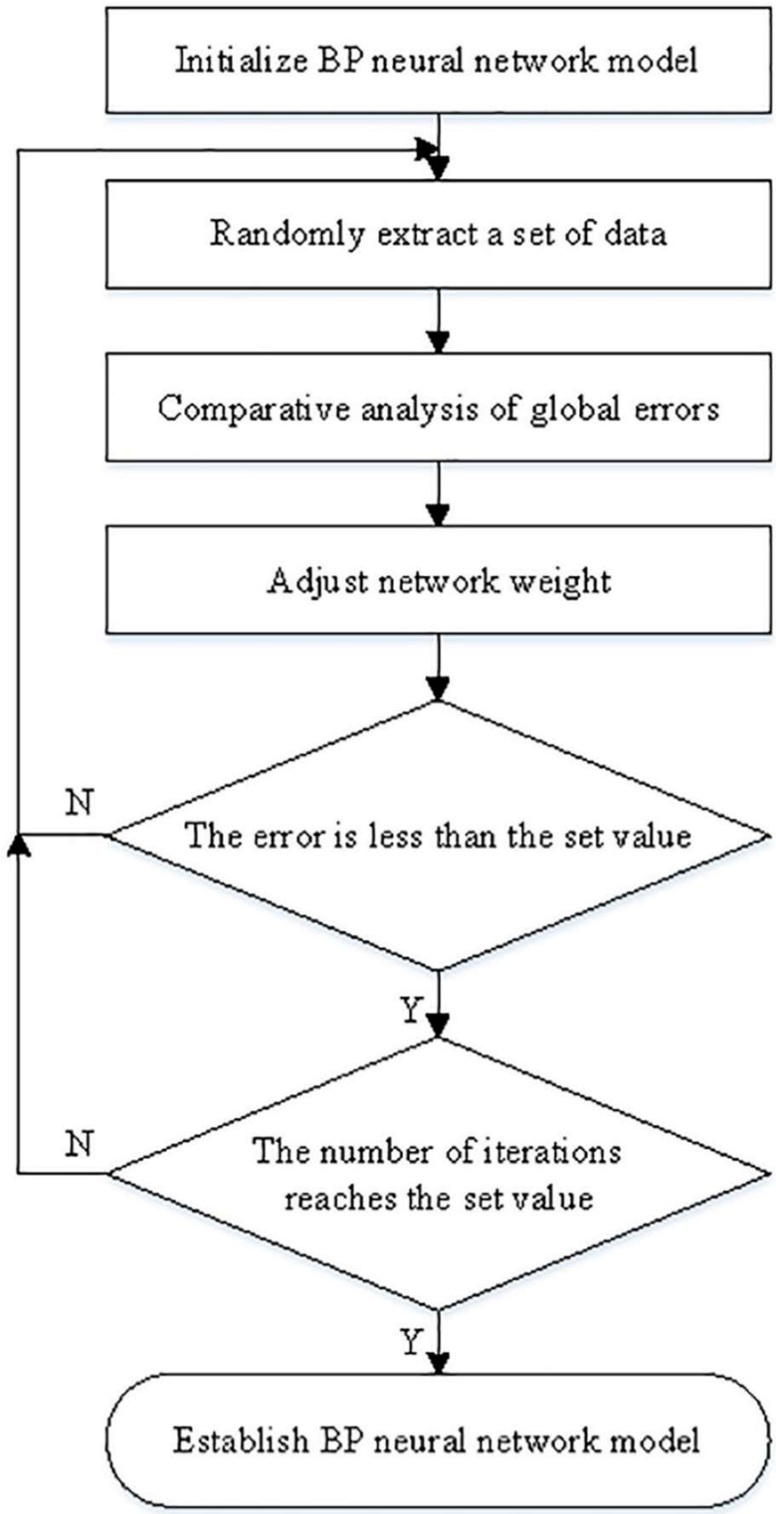
We show that with such sequences, a large artificial gain can be obtained which comes from pattern prediction rather than predicting or compensating the studied channel/phenomena.

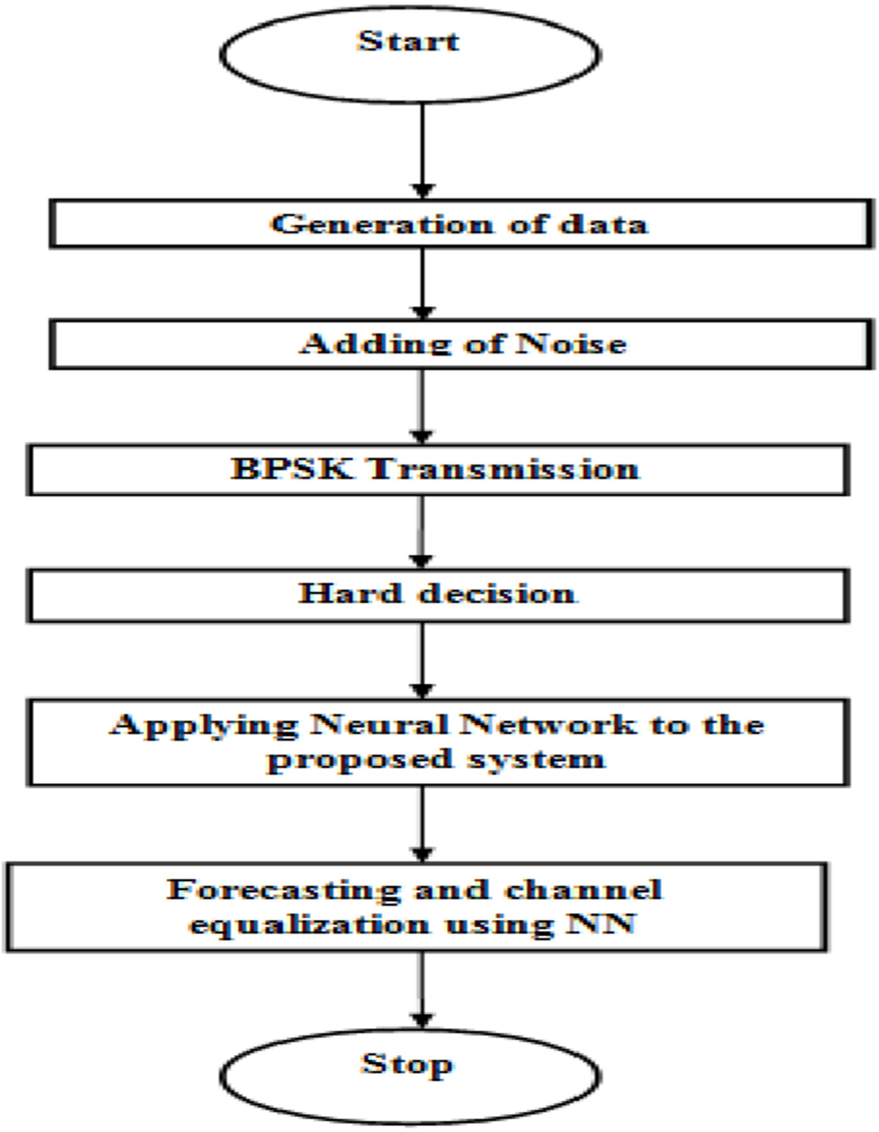
Without this information it is impossible to achieve gain-

- The pattern type that is used.
- The length of the pattern and if it is used (2^n -1)
- The size of the training and evaluation set.

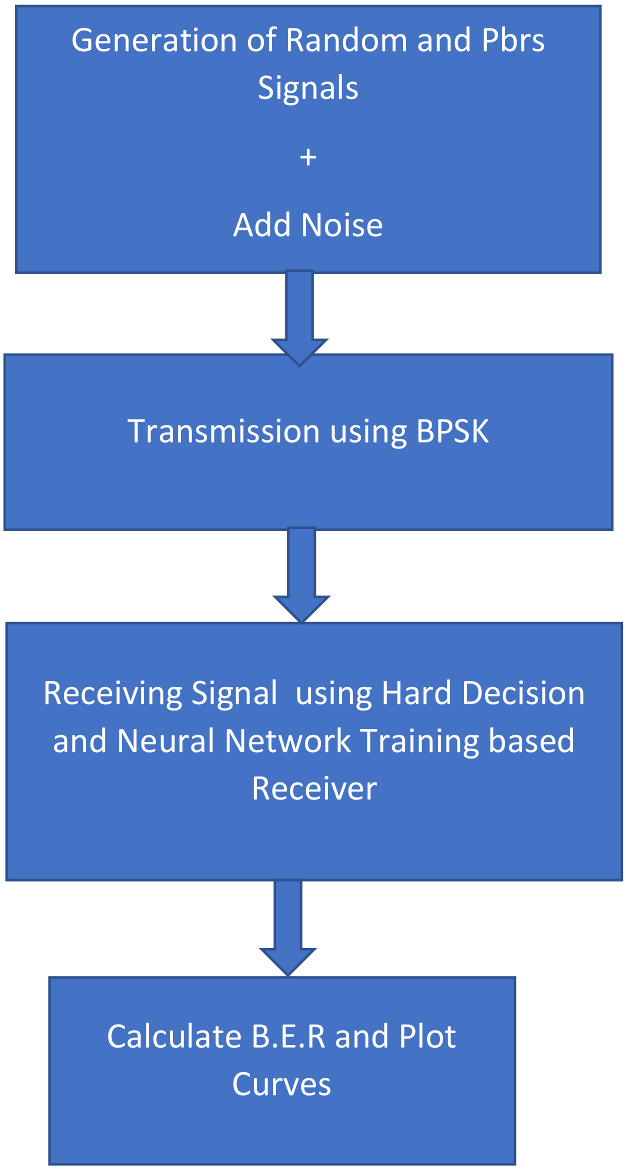
If a different pattern is used for training and evaluation.

This algorithm shows the training done by neural network in the work.


Initialize BP neural network model
Randomly extract a set of data
Comparative analysis of global errors
Adjust network weight
N
The error is less than the set value
Y
N
The number of iterations reaches the set value
Y
Establish BP neural network model


Start
Generation of data
Adding of Noise
BPSK Transmission
Hard decision
Applying Neural Network to the proposed system
Forecasting and channel equalization using NN
Stop

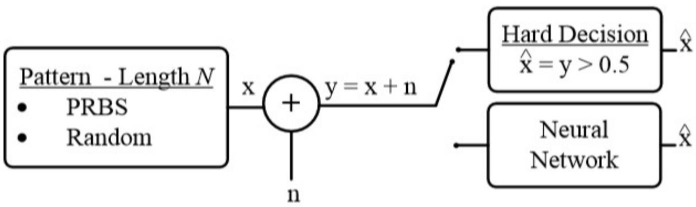
Entire working procedure divided into four major steps:



CHAPTER 4

SYSTEM ARCHITECTURE

To start with, we will investigate a simple scenario, i.e. binary transmission over the additive white Gaussian noise channel (AWGN) as illustrated in Fig. 4.1. We use a simple NN as illustrated in Fig. 4.2. The NN consists of one hidden layer with 8 nodes and conventional rectified-linear and leaky (ReLU) activation functions.



The input layer has L input bits chosen symmetrical around the center bit that is estimated, i.e. L is always an odd number.

For the binary investigation, the output layer consists of two output nodes which corresponds to the probability of a 0 or a 1 being transmitted respectively. For the PAM4 investigation, the output layer has four nodes where the outputs corresponds to the probabilty of each transmitted PAM4 level.

The network is trained using back-propagation with Nesterov's accelerated gradient. This method is similar to stochastic gradient descent but with the gradient taken on the weights with added momentum.

The loss is calculated using multinominal logistic loss. Note that we are not trying to optimize the structure of the NN, the activation functions and the training strategy, but we are rather using a simple structure for demonstration purposes.

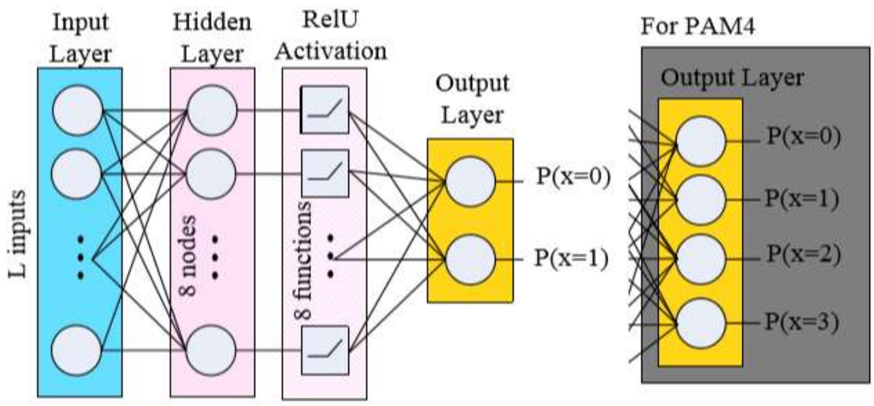


FIG 4.2: SYSTEM ARCHITECTURE

The network is trained from scratch for different input sizes L using PRBS7, PRBS15 or a repeated "random" pattern with length 27. For training, noise is added with a signal-to-noise ratio (SNR) of 10dB. At this point the BER is around 1.3⇥ 102 for hard decision. The length used for the training is 219 blocks of length L. For testing the neural network, we use at least 216 input blocks of either repeated PRBS sequences or instances of a random patterns for each scenario. We always use a new realization of both the noise and the random pattern.

CHAPTER 5
RESULTS AND DISCUSSIONS

SIMULATION RESULTS:

<u>Results</u>

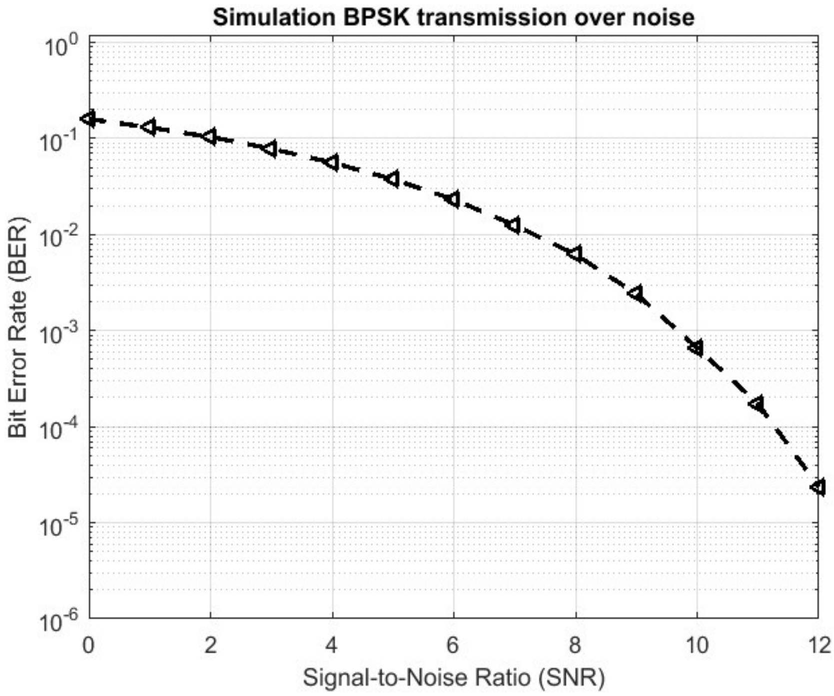


Figure5.1 HD method

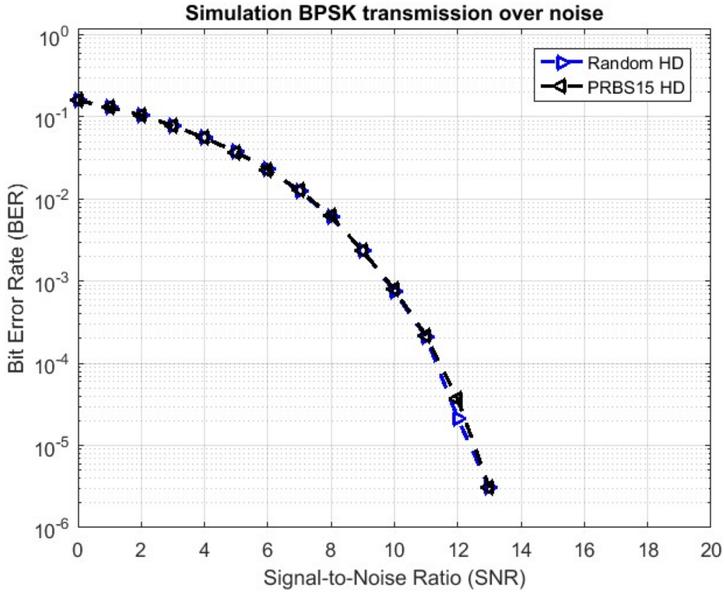


Figure5.2 HD for random and prbs15

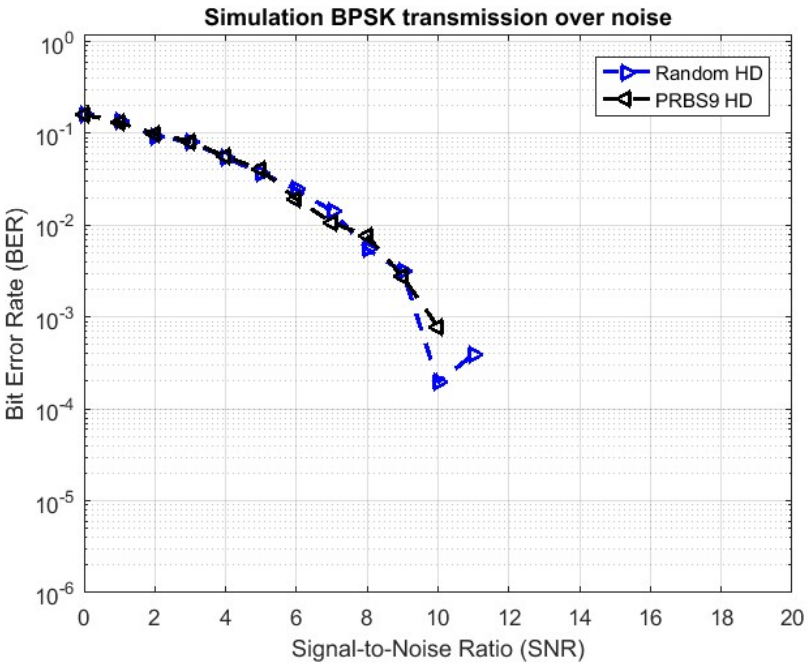


Figure5.3 HD for random and prbs15

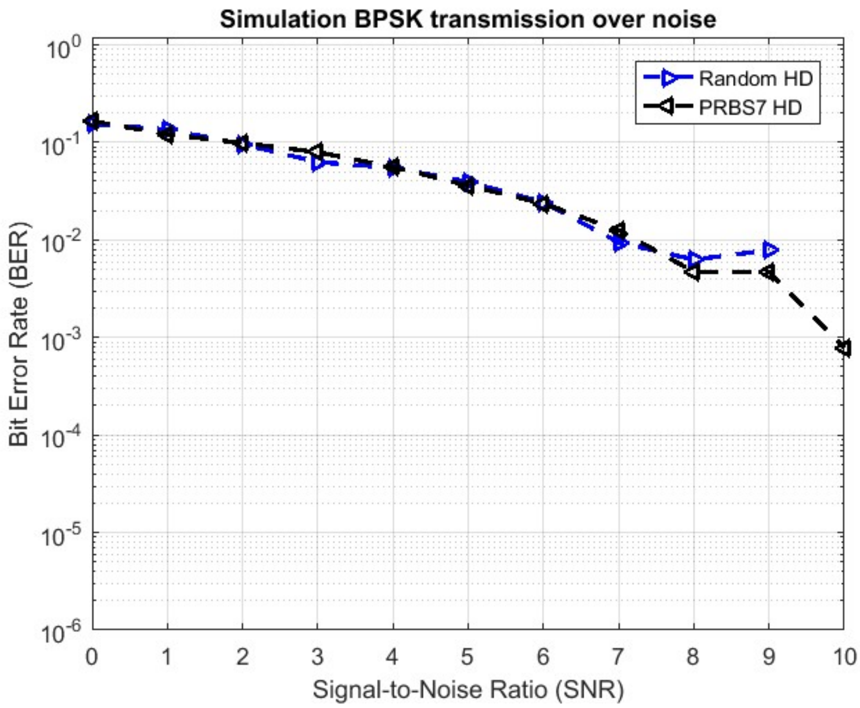


Figure5.4 HD for random and prbs7

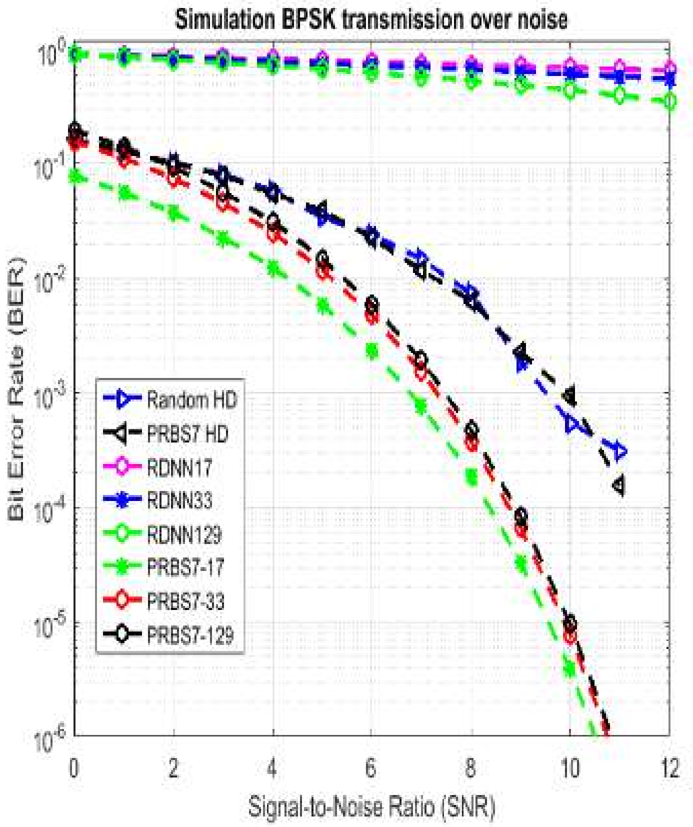


Fig.5.5 Hard decision for Random and PRBS7 pattern ber vs snr curve and for neural network trained on PRBS7 with either PRBS7 or random data is input.

The above figure shows the bervssnr curve using prbs7 and random data as input and use hard decision and neural network. The input length in neural network is 17, 33,129 used in above figure. The first two curves represent the curve of ber vs snr using hard decision. The RDNN17, RDNN33, RDNN129 curve find out by taking random data as input and system is trained using PRBS7data with different neural network input length i.e. 17, 33 and 129.
The ber is 10^{-03} is this setup.

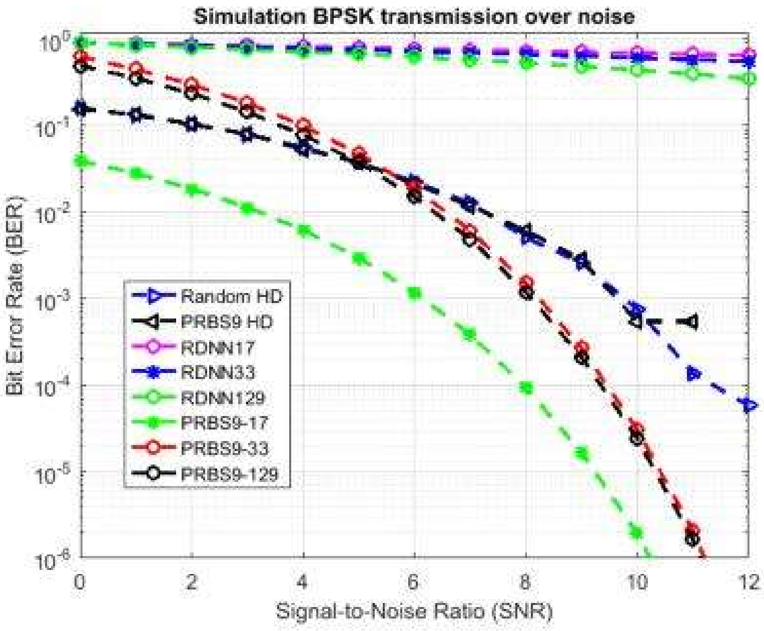


Fig. 5.6 Hard decision for Random and PRBS9 pattern bervssnr curve and for neural network trained on PRBS9with either PRBS9 or random data is input.

The above figure shows the bervssnr curve using prbs9 and random data as input and use hard decision and NN. The input length in neural network is 17, 33,129 used in above figure. The first two curves represent the curve of ber vs snr using hard decision. The RDNN17, RDNN33, RDNN129 curve find out by taking random data as input and system is trained using PRBS9data with different neural network input length i.e. 17, 33 and 129. The ber is 10^{01}
is this setup.

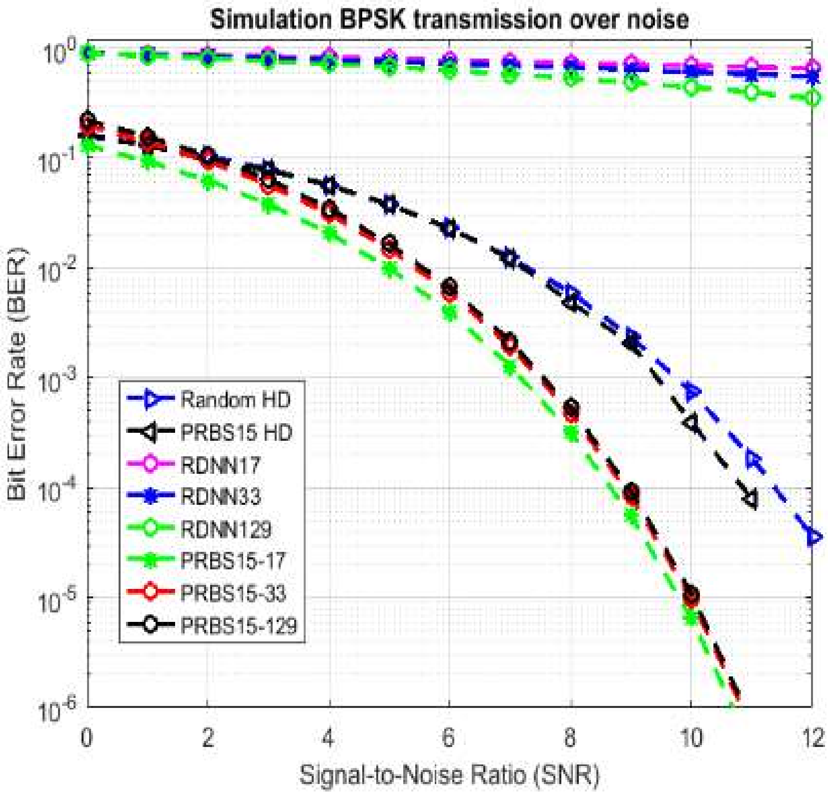


Fig5.7 Hard decision for Random and PRBS15 pattern ber vs snr curve and for neural network trained on PRBS15with eitherPRBS15 or random data is input.

The above figure shows the bervssnr curve using prbs 15 and random data as input and use hard decision and NN. The input length in neural network is 17, 33,129 used in above figure. The first two curves represent the curve of bervssnr using hard decision. The RDNN17, RDNN33, RDNN129 curve find out by taking random data as input and system is trained using PRBS9data with different neural network input length i.e. 17, 33 and 129. The ber is 10^{-03} is this setup.

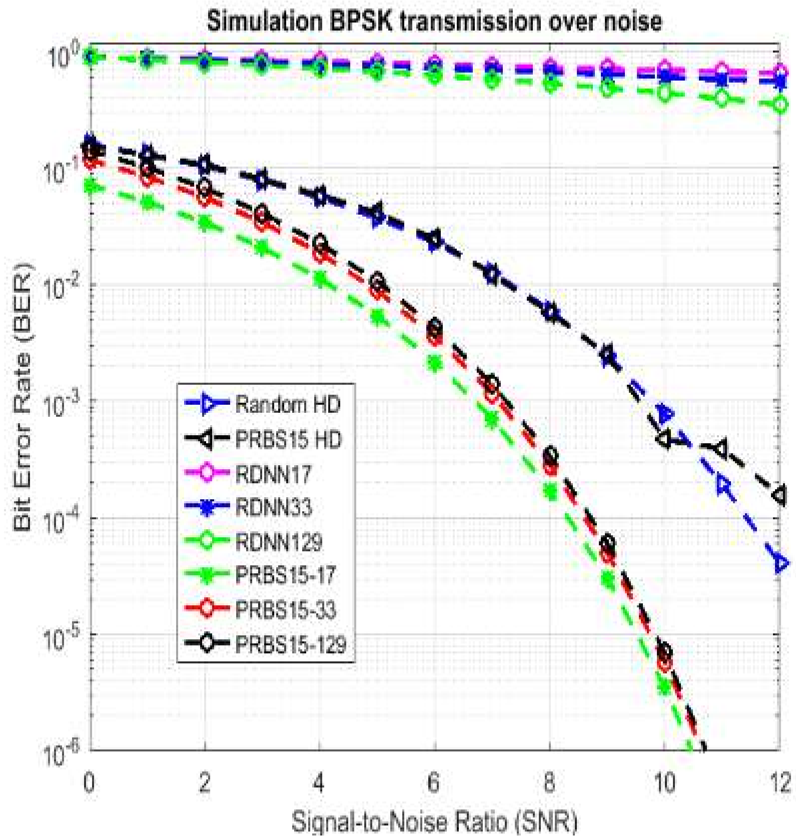


Fig5.8 Hard decision for Random and PRBS15 pattern ber vs snr curve, and for neural network trained on PRBS15 with either PRBS15 or random data is input in this the hidden layer in neural network is 2.

The above figure shows the bervssnr curve using prbs15 and random data as input and use hard decision and NN. The input length in neural network is 17, 33,129 used in above figure. The first two curves represent the curve of ber vs snr using hard decision. The RDNN17, RDNN33, RDNN129 curve find out by taking random data as input and system is trained using PRBS9data with different neural network input length i.e. 17, 33 and 129. The ber is 10^{-04} is this setup.

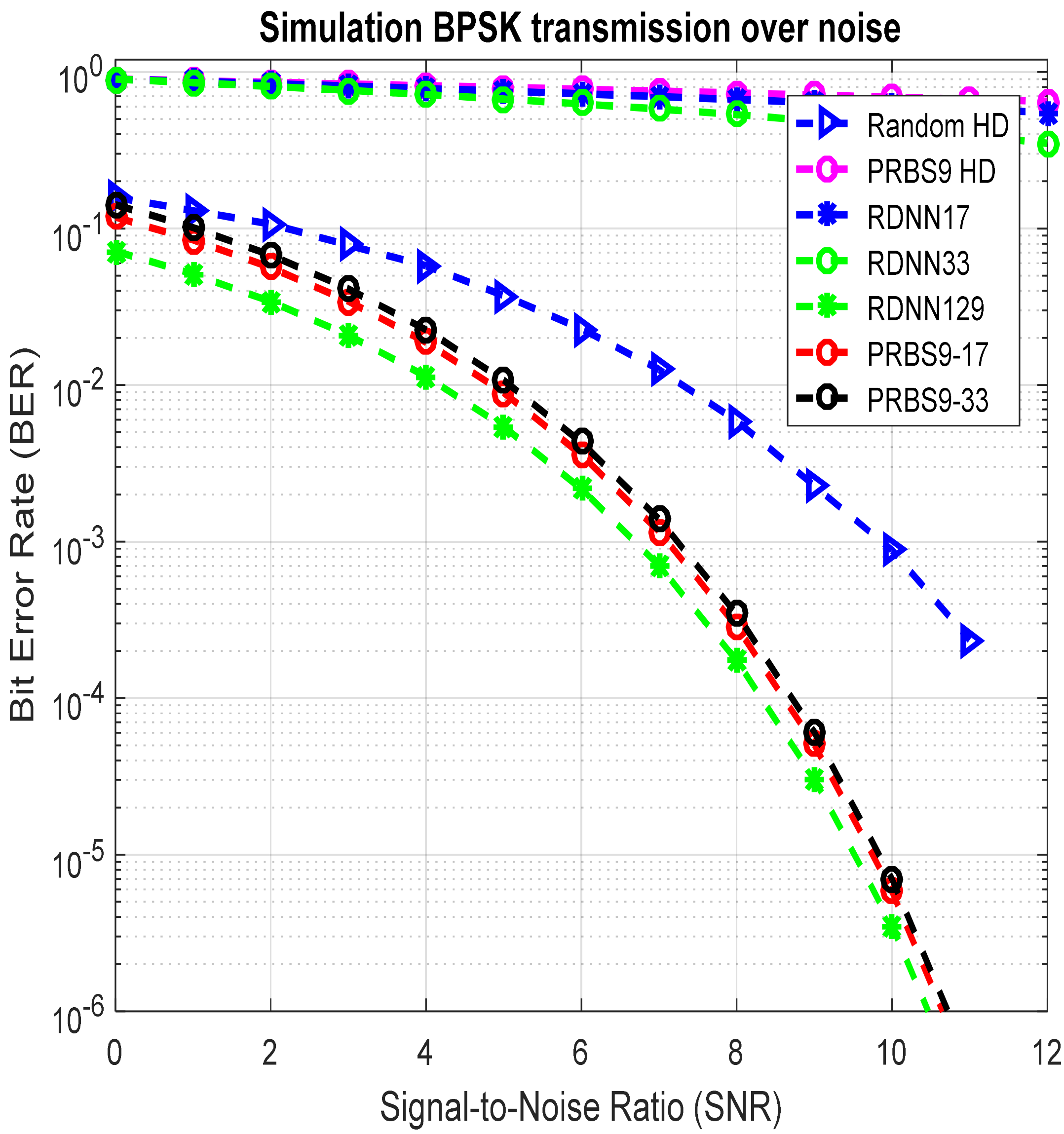


Fig.5.9Hard decision for Random and PRBS9 pattern bervssnr curve and for neural network trained on PRBS9with either PRBS9 or random data is input.

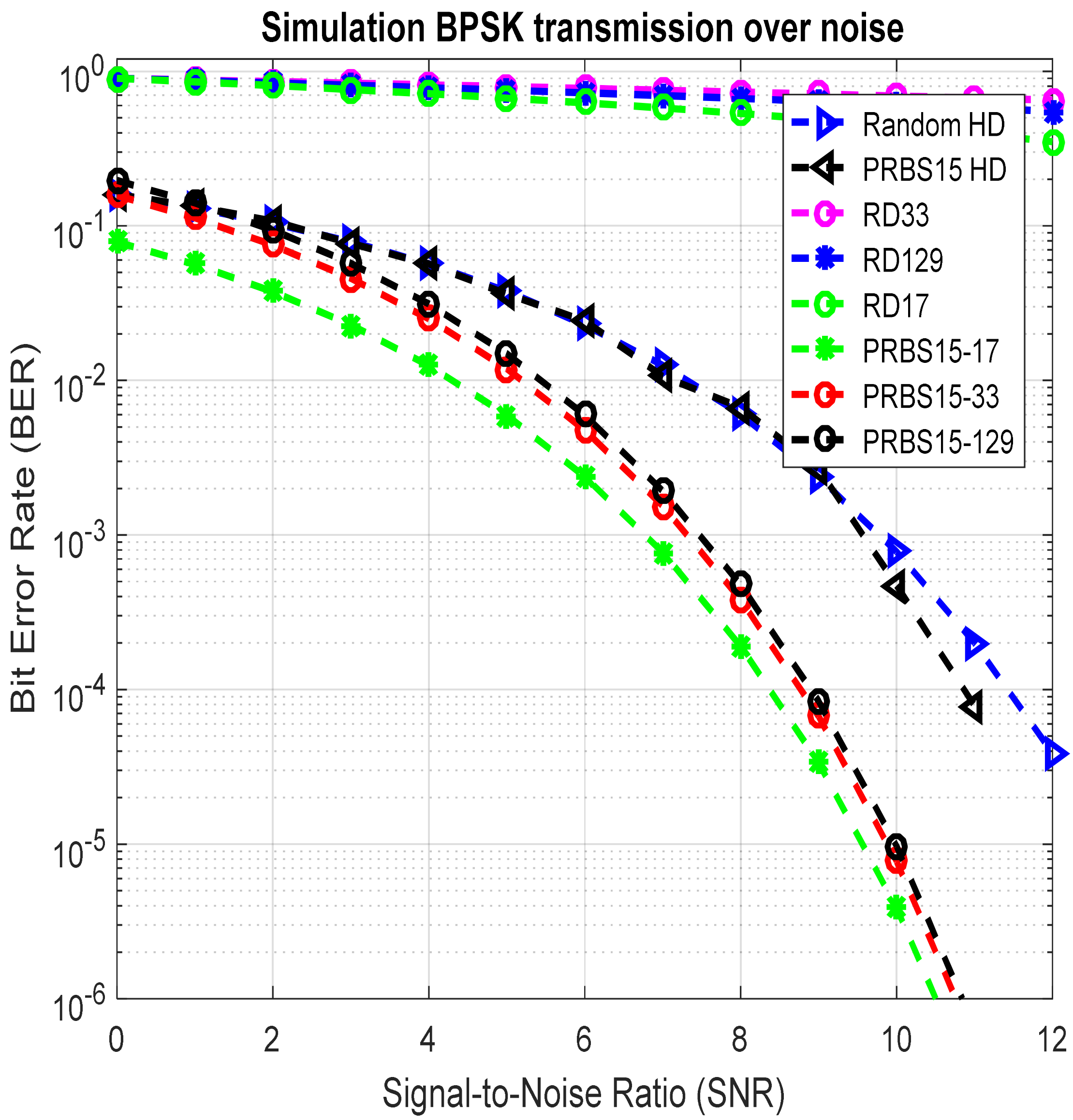


Fig 5.10Hard decision for Random and PRBS15 pattern ber vs snr curve and for neural network trained on PRBS15with eitherPRBS15 or random data is input.

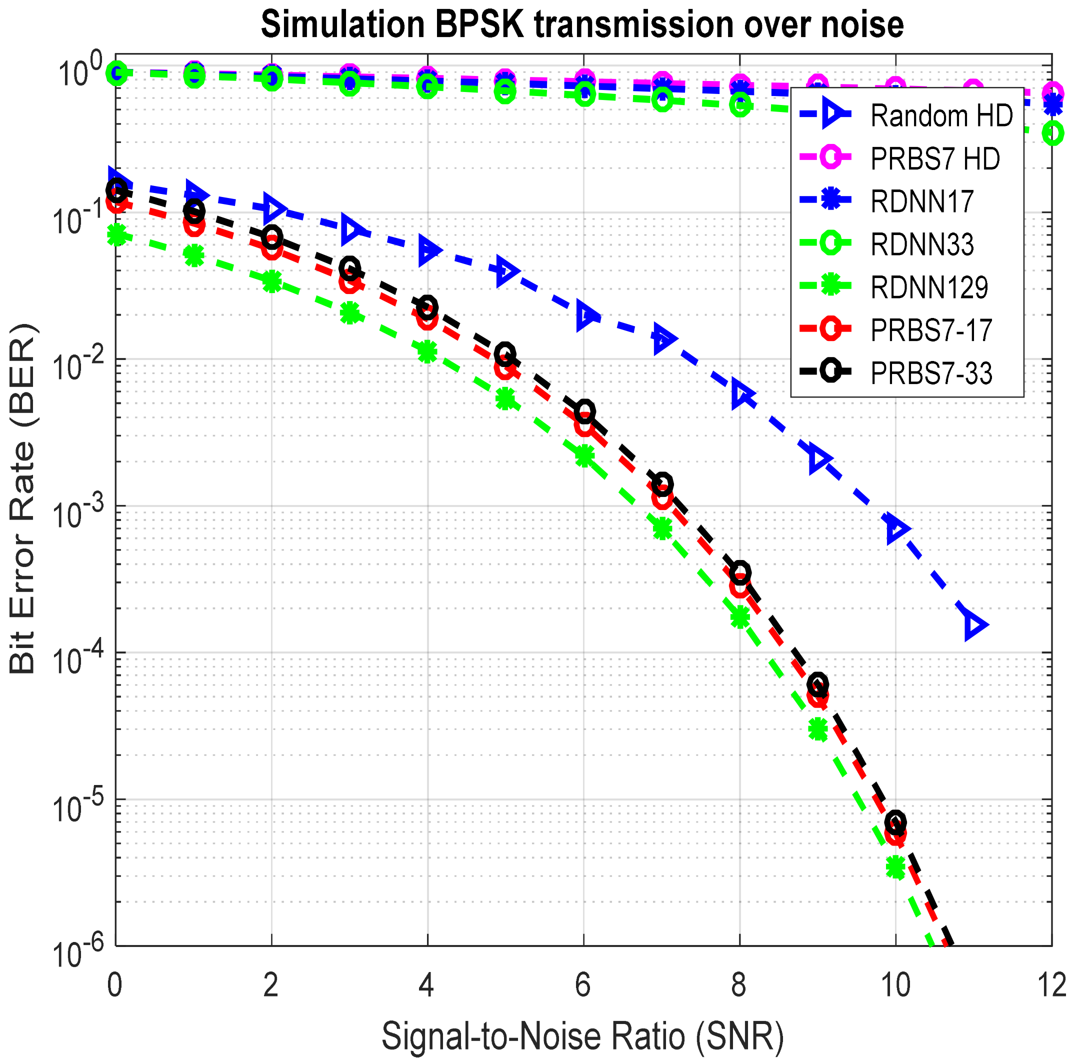


Fig 5.11Hard decision for Random and PRBS7 pattern bervssnr curve and for neural network trained on PRBS7 with either PRBS7 or random data is input.

Conclusion

It has been depicted that the risks of extra evaluating the pursuance profit by the application of neural networks in analytics wherever PRBS series or repetitive small rapid series are characteristically applicable. By the application of neural networks in analytic work, viz. fiber optical communication, here is a requirement to obviously identify the plan for preparation and difficult, counting that prototypes and the span of those which was applied. Devoid of in order, it is unfeasible to arbitrator if the planned system are estimated quite and to moderator if sections or entire of the profit. The DNN is used for predicting instead of ANN which proved to be very efficient as compared to NN.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[2] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” arXiv:1409.0473, 2014.

[3] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learn- ing with neural networks,” in Proc. Advances in Neural Information Processing Systems, 2014, pp. 3104–3112.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Proc. Advances in neural information processing systems, 2012, pp. 1097–1105.

[5] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “Deepface: Closing the gap to human-level performance in face verification,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701–1708.

[6] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath et al., “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 82–97, 2012.

[7] Q. V. Le, “A tutorial on deep learning part 1: Nonlinear classifiers and the backpropagation algorithm,” 2015.

[8] K. Burse, R. N. Yadav, and S. Shrivastava, “Channel equalization using neural networks: A review,” IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, no. 3, pp. 352– 357, 2010.

[9] M. Peng, C. Nikias, and J. G. Proakis, “Adaptive equalization for PAM and QAM signals with neural networks,” in Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems and Computers, 1991, pp. 496–500.

[10] A. Graves, G. Wayne, and I. Danihelka, “Neural turing machines,” arXiv:1410.5401,

2014.

[11] D. Zibar, M. Piels, R. Jones, and C. G. Scha ̈effer, “Machine learning techniques in optical communication,” Journal of Lightwave Technology, vol. 34, no. 6, pp. 1442–1452, 2016.

[12] M. A. Jarajreh, E. Giacoumidis, I. Aldaya, S. T. Le, A. Tsokanos, Z. Ghassemlooy, and N. J. Doran, “Artificial neural network nonlinear equalizer for coherent optical OFDM,” IEEE Photonics Technology Letters, vol. 27, no. 4, pp. 387–390, 2015.

[13] E. Chen, R. Tao, and X. Zhao, “Channel equalization for OFDM system based on the BP neural network,” in Proc. International Conference on Signal Processing, vol. 3, 2006.

[14] E. Giacoumidis, S. T. Le, M. Ghanbarisabagh, M. McCarthy, I. Aldaya, S. Mhatli, M. A. Jarajreh, P. A. Haigh, N. J. Doran, A. D. Ellis, and B. J. Eggleton, “Fiber nonlinearity-induced penalty reduction in CO-OFDM by ANN-based nonlinear equalization,” Optics Letters, vol. 40, no. 21, pp. 5113–5116, 2015.

[15] S. T. Ahmad and K. P. Kumar, “Radial basis function neural network nonlinear equalizer for 16-QAM coherent optical OFDM,” IEEE Pho- tonics Technology Letters, vol. 28, no. 22, pp. 2507–2510, 2016.

[16] J. Estaran, R. Rios-Mu ̈ller, M. Mestre, F. Jorge, H. Mardoyan, A. Kon- czykowska, J.-Y. Dupuy, and S. Bigo, “Artificial neural networks for linear and non-linear impairment mitigation in high-baudrate IM/DD systems,” in Proc. European Conference on Optical Communication (ECOC), 2016, paper M.2.B.2.

[17] S. Warm, C.-A. Bunge, T. Wuth, and K. Petermann, “Electronic disper- sion precompensation with a 10-Gb/s directly modulated laser,” IEEE Photonics Technology Letters, vol. 21, no. 15, pp. 1090–1092, 2009.

[18] S. Gaiarin, X. Pang, O. Ozolins, R. T. Jones, E. P. Da Silva, R. Schatz, U. Westergren, S. Popov, G. Jacobsen, and D. Zibar, “High speed PAM-8 optical interconnects with digital equalization based on neural network,” in Proc. Asia Communications and Photonics Conference, 2016, paper AS1C–1.

[19] D. Wang, M. Zhang, Z. Li, C. Song, M. Fu, J. Li, and X. Chen, “System impairment compensation in coherent optical communications by using a bio-inspired detector based on artificial neural network and genetic algorithm,” Optics Communications, vol. 399, pp. 1–12, 2017.

[20] S. Owaki and M. Nakamura, “Equalization of optical nonlinear waveform distortion using neural-network based digital signal pro- cessing,” in Proc. OptoElectronics and Communications Conference (OECC)/International Conference on Photonics in Switching (PS), 2016, paper WA2-40.

[21] T. S. R. Shen and A. P. T. Lau, “Fiber nonlinearity compensation using extreme learning machine for DSP-based coherent communication systems,” in Proc. Opto-Electronics and Communications Conference (OECC), 2011, pp. 816–817.

[22] F. N. Khan, Y. Zhou, A. P. T. Lau, and C. Lu, “Modulation format iden- tification in heterogeneous fiber-optic networks using artificial neural networks,” Optics Express, vol. 20, no. 11, pp. 12 422–12 431, 2012.

[23] V. Ribeiro, L. Costa, M. Lima, and A. L. Teixeira, “Optical performance monitoring using the novel parametric asynchronous eye diagram,” Optics Express, vol. 20, no. 9, pp. 9851–9861, 2012.

[24] Danshi Wang, Min Zhang”System impairment compensation in coherent optical communication by using a bio inspired detector based on artificial neural network and genetic algorithm” ELSEVIER Optics communications 399(2017)1-12.

[25] Y. LeCun, L. Bottou, G. Orr, and K. R. Mu¨ller. Efficient backprop. In Neural networks: Tricks of the trade. Springer, 1998.

[26] T. Yoshida, T. Sugihara, K. Ishida, and T. Mizuochi, Optical Fiber Communication Conference (OFC) (Optical Society of America, 2014), paper M3C.6.

[27] E. Chen, R. Tao, and X. Zhao, “Channel equalization for OFDM system based on the BP neural network,” in 8th International Conference on Signal Processing (ICSP), Beijing, 2006.

[28] Ajay Gupta, Pankaj Anand”A Survey of Free Optical Space Communication Network Channel over Optical Fiber Cable Communication ”International Journal of Computer Applications(0975-8887)Vol 105-No 10November 2014.

[29] M. Safari and M. Uysal, "Relay-assisted free-space optical communication," in Signals, Systems and Computers, Forty-First Asilomar Conference on, 2007, pp. 1891-1895.

[30] J. C. Juarez, A. Dwivedi, A. R. Mammons, S. D. Jones, V.Weerackody and R. A. Nichols, "Free-space optical communicationsfor next-generation military networks," Communications Magazine,IEEE, vol. 44, 2006, pp. 4651.

[31] C. C. Davis, I. I. Smolyaninov and S. D. Milner, "Flexible optical wireless links and networks," Communications Magazine, IEEE, vol. 41, 2003, pp. 5157.

[32] E. Ciaramella, Y. Arimoto, G. Contestabile, M. Presi, A. D'Errico, V. Guarino and M. Matsumoto, "1.28 Tb/s (32 x 40 Gb/s) WDM transmission over a double-pass free space optical link," in Optical Fiber Communication, Conference on, 2009, pp. 1-3

[33] M. Bass, "Atmospheric optics," in Handbook of Optics, Third Edition, vol. 5, M. Bass, Ed. McGraw-Hill, 2010, pp. 3.3-4.1

[34]J.WangandK.Petermann,"Smallsignalanalysisfordispersiveoptical fiber communication systems," J. Lightw. Technol., vol. 10, no. 1, pp. 96–100, Jan. 1992

[35] G. Katz, D. Sadot, "Radial Basis Function Network Equalizer for Optical Communication OOK System," J. Lightwave Technol. 25, 26312637 (2007).

[36] O. Ozolins et al.,"100 GHz EML for High Speed Optical Interconnect Applications,"Proc. ECOC, M2.4, Dusseldorf (2016).

[37] I. Lyubomirsky, "Machine Learning Equalization Techniques for High Speed PAM4 Fiber Optic Communication Systems," CS229 Final Project Report,Stanford University,2015.

[38] X. Pang et al., "Evaluation of High-Speed EML-based IM/DD links with PAM Modulations and Low-Complexity Equalization," Proc. ECOC,Poster Session,Dusseldorf(2016).

[39] A. Leven and L. Schmalen, “Status and recent advances on forward error correction technologies for lightwave systems,” J. Lightw. Technol., vol. 32, no. 16, pp. 2735–2750, Aug. 2014

[40] K. Onohara et al., “Soft-decision-based forward error correction for 100 Gb/s transport systems,” IEEE J. Sel. Topics Quantum Electron., vol. 16, no. 5, pp. 1258–1267, Sep. 2010.

[41] T.J.Richardson,“Error-floorsofLDPCcodes,”inProc.41stAnnu.Allerton Conf. Commun. Control Comput., Oct. 2003, pp. 1426–1435.

[42] D. A. Morero, M. A. Castrillon, F. A. Ramos, T. A. Goette, O. E. Agazzi, and M. R. Hueda, “Non-concatenated FEC codes for ultra-high speed optical transport networks,” in Proc. IEEE Global Telecommun. Conf., Dec. 2011, pp. 1–5.

[44] M. Li et al., “Low-overhead low-power-consumption LDPC-based FEC solution for next-generation high-speed optical systems,” in Proc. Opt. Fiber Commun. Conf. Exhib., Mar. 2015, pp. 1–3.

[45] L. Schmalen, V. Aref, Junho Cho, D. Suikat, D. Rosener, and A. Leven, “Spatially coupled soft-decision error correction for future lightwave systems,” J. Lightw. Technol., vol. 33, no. 5, pp. 1109–1116, Mar. 2015.

[46] B. S. G. Pillai et al., “End-to-end energy modeling and analysis of longhaul coherent transmission systems,” J. Lightw. Technol., vol. 32, no. 18, pp. 3093–3111, Sep. 2014.

[47] A.Darabiha,A.ChanCarusone,andF.R.Kschischang,“Powerreduction techniques for LDPC decoders,” IEEE J. Solid-State Circuits, vol. 43, no. 8, pp. 1835–1845, Aug. 2008.

[48] T.Mohsenin,D.N.Truong,andB.M.Baas,“Alow-complexitymessagepassing algorithm for reduced routing congestion in LDPC decoders,” IEEE Trans. Circuits Syst. I Reg. Papers, vol. 57, no. 5, pp. 1048–1061, May 2010.

[49] F. Angarita, J. Valls, V. Almenar, and V. Torres, “Reduced-complexity min-sumalgorithmfordecodingLDPCcodeswithlowerror-floor,”IEEE Trans. Circuits Syst. I, vol. 61, no. 7, pp. 2150–2158, Jul. 2014.

[50] E. Cavus, C. L. Haymes, and B. Daneshrad, “Low BER performance estimation of LDPC codes via application of importance sampling to trapping sets,” IEEE Trans. Commun., vol. 57, no. 7, pp. 1886–1888, Jul. 2009.

[51] M. Yang, W. E. Ryan, and Yan Li, “Design of efficiently encodable moderate-lengthhigh-rateirregularLDPCcodes,”IEEETrans.Commun., vol. 52, no. 4, pp. 564–571, Apr. 2004.

[52] Zongwang Li, Lei Chen, Lingqi Zeng, S. Lin, and W. H. Fong, “Efficient encoding of quasi-cyclic low-density parity-check codes,” IEEE Trans. Commun., vol. 54, no. 1, pp. 71–81, Jan. 2006.

[53] Xilinx Inc., 7 Series FPGAs Overview, May 2015, v1.17. [26] G. Marsaglia, “Xorshift RNGs,” J. Statist. Softw., vol. 8, no. 1, pp. 1–6, 2003.

[54] S. Dave, L. Esker, Fan Mo, W. Thesling, J. Keszenheimer, and R. Fuerst, “Soft-decision forward error correction in a 40-nm ASIC for 100-Gbps OTN applications,” in Proc. Opt. Fiber Commun. Conf. Expo. Nat. Fiber Opt. Eng. Conf., Mar. 2011, pp. 1–3.

[55] D.A.Morero,M.A.Castrillon,A.Aguirre,M.R.Hueda,andO.E.Agazzi, “Design tradeoffs and challenges in practical coherent optical transceiver implementations,” J. Lightw. Technol., vol. 34, no. 1, pp. 121–136, Jan. 2016.

[56] Z.Zhang,V.Anantharam,M.J.Wainwright,andB.Nikolic,“Anefficient 10GBASE-T ethernet LDPC decoder design with low error floors,” IEEE J. Solid-State Circuits, vol. 45, no. 4, pp. 843–855, Apr. 2010.

[57] L. M. Zhang and F. R. Kschischang, “Staircase codes with 6% to 33% overhead,” J. Lightw. Technol., vol. 32, no. 10, pp. 1999–2002, May 2014.

[58] C. H¨ager, A. Graell i Amat, H. D. Pfister, A. Alvarado, F. Br¨annstr¨om, and E. Agrell, “On parameter optimization for staircase codes,” in Proc. Opt. Fiber Commun. Conf. Exhib., Mar. 2015, pp. 1–3.

[59] R.Killey,P.Watts,V.Mikhailov,M.Glick,andP.Bayvel,“Electronic dispersion compensation by signal predistortion using a dual-drive Mach–Zehnder modulator,” in Opt. Soc. Amer. Optical Fiber Communication Conf., 2005, Paper OThJ2.

[60]J.McNicol,M.O'Sullivan,K.Roberts,A.Comeau,D.McGhan,andL. Strawczynski,"Electricaldomaincompensationofopticaldispersion," in Opt. Soc. Amer. Optical Fiber Communication Conf., 2005, Paper OThJ3.

[61] J.WangandK.Petermann,"Smallsignalanalysisfordispersiveoptical fiber communication systems," J. Lightw. Technol., vol. 10, no. 1, pp. 96–100, Jan. 1992.

[62] K. Petermann, Laser Diode Modulation and Noise, T. Okoshi, Ed. Norwell, MA: Kluwer, 1988.

[63] S. Haykin, Neural Networks. A Comprehensive Foundation, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[64] X. Hu, Y. Shi, and R. Eberhart, "Recent advances in particle swarm," in Congress on Evolutionary Computation, 2004 (CEC2004) , vol. 1, pp. 90–97.

[65] S. Warm, C. A. Bunge, T. Wuth, and K. Petermann, "Electronic dispersion precompensation using a directly modulated laser," in ECOC

[66] K. Jarrett, K. Kavukcuoglu, M. A. Ranzato, and Y. LeCun. What is the best multi-stage architecture for object recognition? In International Conference on Computer Vision, pages 2146–2153. IEEE, 2009.

[67] A. Krizhevsky. Learning multiple layers of features from tiny images. Master's thesis, Department of Computer Science, University of Toronto, 2009.

[68] A. Krizhevsky. Convolutional deep belief networks on cifar-10. Unpublished manuscript, 2010.

[69] A. Krizhevsky and G.E. Hinton. Using very deep autoencoders for content-based image retrieval. In ESANN, 2011.

[70] Y. Le Cun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, L.D. Jackel, et al. Hand written digit recognition with a back-propagation network. In Advances in neural information processing systems, 1990.

[71] Y. LeCun, F.J. Huang, and L. Bottou. Learning methods for generic object recognition with invariance to pose and lighting. In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, volume 2, pages II–97. IEEE, 2004.

[72] Y. LeCun, K. Kavukcuoglu, and C. Farabet. Convolutional networks and applications in vision. In Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, pages 253–256. IEEE, 2010.

[73] H. Lee, R. Grosse, R. Ranganath, and A.Y. Ng. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In Proceeding of the26th Annual International Conference on Machine Learning, pages 609–616. ACM, 2009.

[74] T. Mensink, J. Verbeek, F. Perronnin, and G. Csurka. Metric Learning for Large Scale Image Classification: Generalizing to New Classes at Near-Zero Cost. In ECCV - European Conference on Computer Vision, Florence, Italy, October 2012.

[75] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In Proc. 27th International Conference on Machine Learning, 2010.

[76] N. Pinto, D.D. Cox, and J.J. DiCarlo. Why is real-world visual object recognition hard? PLoS computational biology, 4(1):e27, 2008.

[77] N. Pinto, D. Doukhan, J.J. DiCarlo, and D.D. Cox. A high-throughput screening approach to discovering good forms of biologically inspired visual representation. PLoS computational biology, 5(11):e1000579,

[78] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber, "Deep, big, simple neural nets for handwritten digit recognition," Neural Comput., vol. 22, no. 12, pp. 3207–3220, 2010.

[79] G. E. Hinton and R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” Science, vol. 313, no. 5786, pp. 504–507, 2006.

[80] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, “An empirical evaluation of deep architectures on problems with many factors of variation,” in Proc. 24th Int. Conf. Machine Learning, 2007, pp. 473–480.

[82] J. Pearl, Probabilistic Inference in Intelligent Systems: Networks of Plausible Inference. San Mateo, CA: Morgan Kaufmann, 1988.

[83] G. E. Hinton, “Training products of experts by minimizing contrastive divergence,” Neural Comput., vol. 14, pp. 1771–1800, 2002.

[84] G. E. Hinton, “A practical guide to training restricted Boltzmann machines,” Tech. Rep. UTML TR 2010-003, Dept. Comput. Sci., Univ. Toronto, 2010.

[85] G. E. Hinton, S. Osindero, and Y. Teh, “A fast learning algorithm for deep belief nets,” Neural Comput., vol. 18, no. 7, pp. 1527–1554, 2006.

[86] T. N. Sainath, B. Ramabhadran, and M. Picheny, “An exploration of large vocabulary tools for small vocabulary phonetic recognition,” in Proc. IEEE Automatic Speech Recognition and Understanding Workshop, 2009, pp. 359–364.

[87] A. Mohamed, T. N. Sainath, G. E. Dahl, B. Ramabhadran, G. E. Hinton, and M. Picheny, “Deep belief networks using discriminative features for phone recognition,” in Proc. ICASSP, 2011, pp. 5060–5063.

[88] A. Mohamed, G. Hinton, and G. Penn, “Understanding how deep belief networks perform acoustic modelling,” in Proc. ICASSP, 2012, pp. 4273–4276.

[89] Y. Hifny and S. Renals, “Speech recognition using augmented conditional random fields,” IEEE Trans. Audio Speech Lang. Processing, vol. 17, no. 2, pp. 354–365, 2009.

[90] A. Halberstadt and J. Glass, “Heterogeneous measurements and multiple classifiers for speech recognition,” in Proc. ICSLP, 1998.

[91] A. Mohamed, D. Yu, and L. Deng, “Investigation of full-sequence training of deep belief networks for speech recognition,” in Proc. Interspeech, 2010, pp. 2846–2849.

[92] T. N. Sainath, B. Ramabhadran, M. Picheny, D. Nahamoo, and D. Kanevsky, “Exemplar-based sparse representation features: From TIMIT to LVCSR,” IEEE Trans. Audio Speech Lang. Processing, vol. 19, no. 8, pp. 2598–2613, Nov. 2011.

[93]. Tompson, J., Goroshin, R. R., Jain, A., LeCun, Y. Y. & Bregler, C. C. Efficient object localization using convolutional networks. In Proc. Conference on Computer Vision and Pattern Recognition http://arxiv.org/abs/1411.4280 (2014).

[94]. Taigman, Y., Yang, M., Ranzato, M. & Wolf, L. Deepface: closing the gap to human-level performance in face verification. In Proc. Conference on Computer Vision and Pattern Recognition 1701–1708 (2014).

[95] Hadsell, R. et al. Learning long-range vision for autonomous off-road driving. J. Field Robot. 26, 120–144 (2009).

[96] Farabet, C., Couprie, C., Najman, L. & LeCun, Y. Scene parsing with multiscale feature learning, purity trees, and optimal covers. In Proc. International Conference on Machine Learning http://arxiv.org/abs/1202.2160 (2012).

[97] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. J. Machine Learning Res. 15, 1929–1958 (2014).

[98]. Sermanet, P. et al. Overfeat: integrated recognition, localization and detection using convolutional networks. In Proc. International Conference on Learning Representations http://arxiv.org/abs/1312.6229 (2014