Cortical Learning Algorithm Implemented as HTM Deep Learning Model

Hamid Masood Khan[1], Sheikh Muhammad Saqib[1], Jamal Abdul Nasir[1], Muhammad Asif Nadeem[2], Arslan Naveed[2]

[1]Institute of Computing and Information Technology, Gomal University, Dera Ismail Khan, KPK, Pakistan.
[2]M.Phil. Scholars Physical Education

Emailof Author#1:hamidmasoodkhan@hotmail.com, Email of Author#2: saqibsheikh4@gu.edu.pk, Email of Author#3: jamal_icit@hotmail.com

Abstract

A proposed theoretical framework for sequence learning, Hierarchical Temporal Memory (HTM) model has been developed upon the working principles of human neo-cortex.In HTM model, all the elements of neo-cortex like neurons and its cell body, axon, dendrites and synapse etc. have been restructured by assigning different names like cells, regions, levels, grid cells, mini columns and duties like active state, bottomup, feedforward, hierarchy, density, permanence etc., on same grounds for information processing and building a structural image of the surrounding environment. As there are different types of inputs to human brain from different sensors, similarly the HTM model can have different types of inputs. Functions of all the six layers identified in neo-cortex are copied and modeled in HTM theory for information processing but making it different from existing neural network processing structure by introducing Sparse Distributed Representations as that of human brain.Like human brain, HTM can handlealltypes of data.HTM gave birth to Semantic Folding Theory (SFT)to represent a data stream for processing in the form of Sparse Distributed Representation (SDR). For natural language perception and production, SFTcarriesa solid structural backgroundto the fundamentals of semantic foundation during the phase of language acquisition.

Key words:

Hierarchical Temporal Memory, semantic folding theory, neocortex, sparse distributed representations.

1. Introduction:

Neuroscientists in their evolutionary research claimed that our brain is a memory system that learns from everyday life's experiences of surrounding environment. These learned experiences aresequentially stored as structural representation of events and their association with each other. Brain initiates future predictions based on these structural representations of surrounding events. This prediction system leads towards creativity and perception which are the foundations of human intelligence. [1-4].Human brain identifies these coming inputs and predicts the next organized temporal sequences to each and every function of that inputs to brain, includingactive observable perception,speech recognitionand natural vision. Inference from the behaviour of a person in specific situationswas a challenge [5]. However, human behaviour's, prediction, perception, intention andbeliefs depend upon, his own mental capabilities,reasoning, inference and social interactions[6]. According to Neuroimaging studies, during processing,multiple cortical regions are involved [7, 8].Human cortex continuously processes streams of data and information coming from sensorslike retina, cochlear, somatic and constructs a dynamic rich and general spatio temporal model of surrounding environment[9, 10]. Survival in natural environments is dependent on sensory inputs in the form of temporal sequences, received, recognized and predicted [11]. The recent known properties of neurons in cortex lead to the development of a theoretical framework termed as Hierarchical Temporal Memory (HTM) model for sequence learning [1, 2, 12].

The HTM model has the capability to achieve comparable accuracy to the most common and developed algorithms like ARIMA, EGADS etc.[13]. All the important properties of sequence learning and continuous online learning are enclosed in this model.HTM model can resolvenumber of predictionssimultaneously. This Model satisfactorily performs better in complex multifacedsequences with sensor noise andfaults [14]. HTM sequence memory algorithms increase our understanding about “how the brain solves the problem of sequence learning?”. HTM algorithms are alsoappropriate to learn about issues of sequence learning from online data streams.HTM uses adifferent approachtowards the solution of the problems rather than conventional functional mapping [5, 15]. Neuro researchers have developed more sophisticated models for three dimensional structure of the cortical neural network[16]. This micro level research upon internal circuitrary and functions of dendrites, axons and synapses enabled the formation of HTM model keeping forth all the functional parameters of human brain. Similarly interconnection in cortical tissues and functional paths identification lead to the sophisticated algorithmic changes having ultra-advance imaging techniques.[17]. Complete knowledge and functionality of all prevailing pathways, the scientists have prediction to the general information architecture of the brain[18]. In contrast to all existing data-driven methods and models, HTM theory converges towards understanding and identifying principles as well asfunctional mechanisms through which the human neo-cortex operates[19-22]. Behavioural research is matched with identified characteristic of human neo-cortex. All aboveHawkin’s cortical learning principles explained were theoretically implemented in HTM theory.

Biologically convenient and easy data configuration are the deriving forces towards the formulation of Sparse Distributed Representation to process data upon the principles of cortical information processing, in real time environments [23, 24].The sparse distributed representations, an input model developed for, and used by HTM model is capable of handling temporal and spatial sequences and maintain various predictions till any event becomes clear [25]. In today’s world, drastic evolution of time dependent real time data streams, the detection of anomalous behavior gives significant information in critical situations. Detection of anomalous behaviour,in continuous data streams is a more complicatedprocess for sensors to detect and predict after learning is made easy using HTM theory[26].Semantic Folding Theory (SFT) is developed upon the working principles of Hierarchical Temporal Memory Theory associated to biological data and understanding, of processing principles of neo-cortices. This theory was developed in a way that maximum features and characteristics of human language understanding in neo-cortex are accommodated in it[25]. Semantic information manipulation in neo-cortex, when significantlyunderstood by the neuroscientists, lead to the evolution of Semantic Folding Theory for natural languages production and perception[23]. This latest innovative approach solved the representational problems by incorporating the words context in way that is accessible to machine.Language information processing down to its meanings level empower a better understanding[19-22].

1.1 Expected Audience:

This study is very much informative for researchers working in the field of opinion mining (also called sentiment analysis) using traditional methods for tokenization, lemmatization, stemming, stop word removal and lexical analysis. The pre-processing speeds up by using HTM semantic folding theory. Hawkin's theory based upon human language learning model is a new opening in the field of natural language processing[23, 27-29]. The texts could be easily processed using their semantic meanings. All research carried out on document classification, text mining, sentiment analysis and those working on anomaly detection are the competent audience of this article.

1.2 Methodology and Scope:

Under discussion review is based upon the postulates of HTM theory as proposed by Jeff Hawkins. The literature contained in this article is impartial, all available, complete, comprehensive which is collected from the research papers and videos available at www.http://numenta.com, lecture slides of founder of HTM theory Jeff Hawkins, available at www.http://slideshare.com, and discussion hour videos of Numenta community, keeping focus on HTM and its functionality. A dire need was felt to collect all relevant information of this theory to facilitate the researchers of natural language processing and sentiment analysis. It is ensured that the information and methods be captured to facilitate the researchers without any inclination and biasness by taking all the published available literature. The delivered lectures and "Numenta hour" discussions about HTM theory were also pen captured and incorporated appropriately.

2. Theoretical Information Extraction

In this article the historic evolution of Artificial intelligence that lead towards HTM in the following way as shown in fig 2.1:

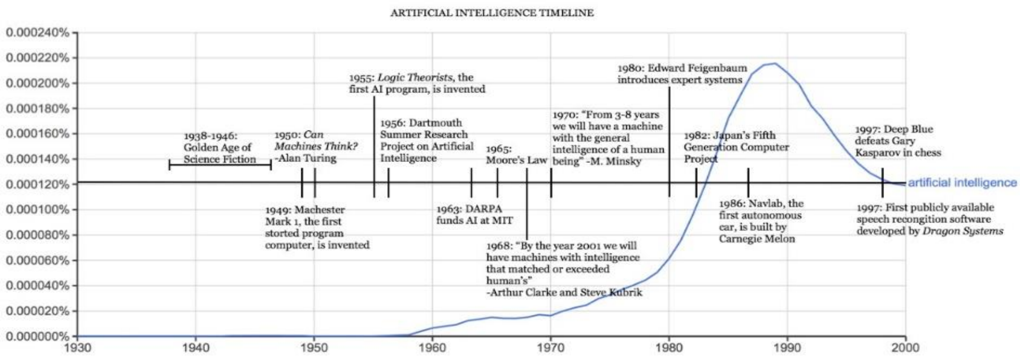

Fig 2.1: Artificial Intelligence timeline.

2.1 Neural Network:

A complex system of neurons in humanbrain is called neural network. Based upon working principles of neural network, artificial neural network is composed of artificial neurons or nodes for solving AI problems [30]. These are computational models developed on the known features of biological neural networks, used to formulate and calculate unknown functions, inspired from the neuronsbehaviour and their input and output communication signals.The semantics of communication structure of neuronsis an area of HTM research[31].Artificial neural networks are based upon variety of deep learning technologies[32].Neural networks also known as software-based computer models, with number of complex learning algorithms and topologies[33]. Neurons are represented as hardware components in these artificial neural networks[34].

2.2Machine learning:

Enabling computers to learn and act intelligently, and autonomously expand their learning with passage of time through input data, information and observation is the basic definition of machine learning[35]. Machine learning has become universal and indispensable for solving complex problems in most sciences.Machine learning enables computers to improve automatically through experience. Machine learning is most rapidly growing field that encircles computer science, statistics, artificial intelligence and data science. Recent advancements in machine learning lead to the new learning algorithms and concepts through online data and low-cost computations. Informationdependent machine learning approachesinitiated in technology and commerce lead to the more evidence-based decision making in manufacturing, education, health care, financial modeling and marketing [36].

Machinelearning research has been achieving great advancements in number of fields. The following four enlisted related directions are: [36]

- Classification perfection accuracy by improving classifiers learning capability.
- Supervised learning algorithms improvement methods,
- Reinforcement learning,
- Stochastic models learning improvement.

Supervised learning, either for analysis of data sets, or as a sub goal of a more complex problem[37], in the form of regression (for continuous type of outputs) and classification (for discrete type of outputs) is an important constituent of statistics and machine learning [38].

2.3 Hierarchical Temporal Memory:

Upon the understanding principles and functional capabilities and proficiencies of neo-cortex, a theoretical HTM framework has been developed. Hierarchical Temporal Memory (HTM) is a latest emerging machine learning technology based on structural and learning algorithmic properties of the neo-cortex. All the human actions are performed by the neo-cortex including vision, hearing, touch, movement, language (learning and speaking) and planning. Such diversity need specialized neural algorithm. But on the contrarythe neo-cortex displays a uniform pattern of neural circuitrythatenforces a shared set of algorithms to accomplishvarioustasks and functions.

HTM learning algorithm implemented a sufficient number of subsets of neo-cortex functions. Programming HTMs is not just like programming traditional computers. HTMs are programmed and trained through data coming from different sensors.The capabilities of HTMs are determined on a large scale by exposing them to different types of data streams. HTMs are type of neural network.HTM model neurons (called cells), are arranged in columns, layers,regions, and has hierarchy.Basically,HTM is developed as a memory-based learning system. HTMs are logically trained upon varying data, andtheir dependency is on storage of extended binary vectors in the form of SDRs. Classic computer memory does not have any time constraint to observe.In flat computer memory, a programmer has opportunity to implementdifferent types of data structures and data organizations with a maximum control of how and where to store information.But, HTM memory has restrictions of hierarchy and time thus information is always stored in a distributed way. The size and hierarchy of HTM is specified by the programmer or user, but the complete control of information storage remains with HTM. General purpose computers are used to model the HTM networks with their own computing constraints like unique parameter of hierarchy, time and sparse distributed representations.HTM's performance in multi informational streams is also examplary. Hence HTMs are learning machines with a main feature of prediction, that could be exposed to number of different forms of problems.

Hierarchy and regions are the fundamental components of an HTM network. Characteristically, each level in hierarchy represents an HTM region, ascending inhierarchy leads towards convergence. By the convergence of several elements in a child region formation of an element in a parent region is done. However, descending down the hierarchy diverges the information, due to feedback connections.Multiple HTM networks can also be combined when the input data have number of sources or senses. Each network has its own convergence towards top. Learning efficiency is main advantage of hierarchical organization by reducing training time and memory usage combing atupper level[39].

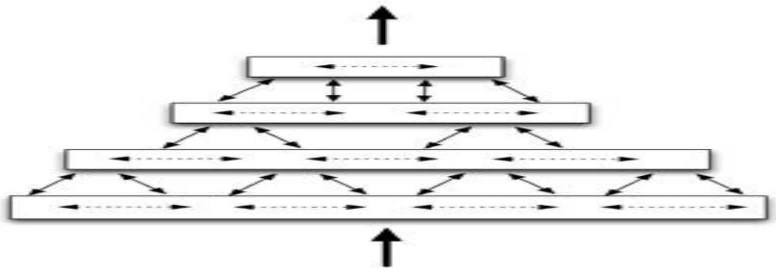

Fig2.2: Diagram of four HTM regions in a four-level hierarchy[29].

Generalization of expected behaviour is done by Sharing representations in a hierarchy. For example, when someone sees a new animal with mouth and teeth, it will be easily predicted that the animal uses mouth for eating and uses teeth for biting purpose. Hierarchy in HTM's is the main factor of inheritance of known properties of its subcomponents for a new object.A tradeoff matrix is required for the allocation of memory at each level and for the number of levels required. [40].

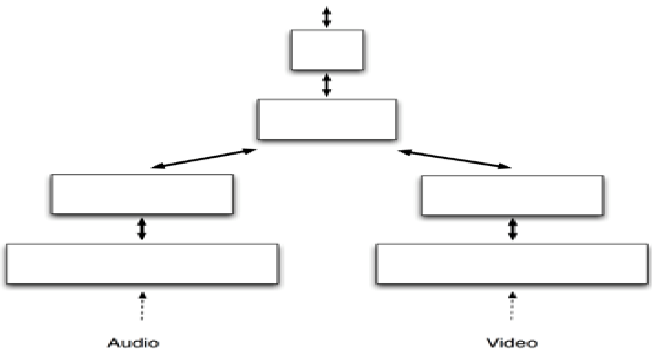

Fig 2.2 (a): Converging networks from different sensors[29]

Depending upon the input statistics and allocatedmemory resources, HTMsinevitably learn the appropriateillustrations at each level. A level will form larger and complex representations when more memory is allocated ultimately reducing the number of levels, means fever but necessary hierarchical levels. If less memory is allocated, representations on a level would be smaller and simpler.For spatial hierarchy problemshaving temporal statistics, the temporal learning algorithms in HTMs are perfect. In brief, generalized formation, reduced training time and reduced memory usage are the advantages of hierarchies. However, many problems of simpler prediction nature could be answered with a single HTM region[21, 41].

2.4 Cortical facts:

It is an organ of memory that learns through sensory organs like retina, cochlear and somatic. They form identical patterns of actions on cortex[42-44].The neo cortex learns predictive model from fast changing sensory data. This predictive model generates predictions, anomalies and action (behaviour). Most of the sensory changes are due to movement in sensory organs. The neo cortex learns a Sensory-Motor-Model from around world. Fig 2.3 shows the facts described as follow:

1. Human Neo-Cortex is a thin 2.5mm sheet of cells {containing 60 billion neurons) which is remarkably uniform functionally and anatomically.
2. Hierarchy of identical regions in human cortex is larger than other mammals.
3. The next are the six cellular layers containing mini columns each. These mini columns contain neuron of two types.[18, 40, 45]. Each Neuron have w/1000’s synapses, in which 10% are proximal and 90% are distal. Active distal dendrites polarize the cell[46-48]. Continuous changing of synaptic weights in brains learning is known as synaptogenesis.
4. In layers each region learns sequences. Stability in learning increases going up hierarchy, if input are predictable[46].

Cortical Facts

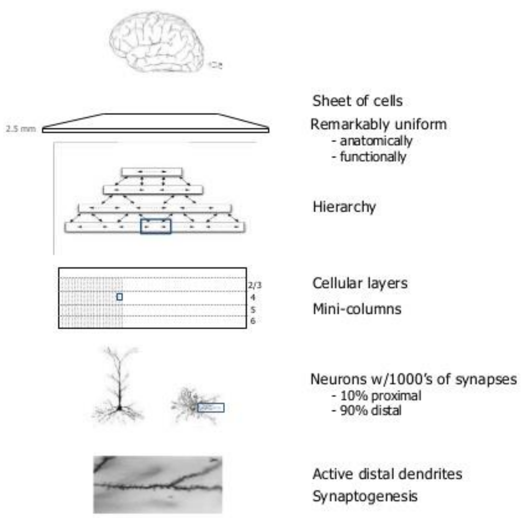

Fig 2.3 cortical facts (http://www.numenta.com)

Vernon Mountcastle had proposed that all neo-cortex regions are primarily the same. There is no difference of internal circuit and their functions. Different types of inputs are given to the different regions.Replication occurs in a small volume of cortex. This small unit of cortex is called cortical column[49]. Accordingly recent experiments proved that, neo-cortex has grid cells [50]. Hawkins proposed that object centric location with in grid cells sensory region is also present, that activates on sensory input by the neo-cortex [4].Hence, in each column of neo-cortex, neurons like grid cell exist. Properties of Grid cells are unique having a membrane [51].

According to Stensola, neuron that convert into activestate at number of locations in different environments is called grid cell. Within an environment all active grid cells at different locations form a module called grid cell module[52].From these cortical facts Hawkinsderived that:

- Every cortical column has grid cells.
- Active grid cells represent a location.
- Location is updated with trackintegration in to grid cells.
- Cortical column learn model of objects while grid cells learn model of environment. [4]

From above it is clear that model of objects is learned by each cortical column completely. This is achieved by location grid cell plus integrating input, and then integrating movements[53].

2.5Hierarchical Temporal Memory(HTM) Model Theory:

HTM is an abbreviation for Hierarchical Temporal Memory, a term used todefineworking models of neo-cortex [25].Developed upon the operational and algorithmic designing of the neo-cortexHierarchical Temporal Memory (HTM) is a machine learning technology. Having so many diverse sets of cognitive functions, it is expectedfrom neo-cortex to implement likewise diverse suite of dedicated neural algorithms. But on the contrary neo-cortex exhibits a remarkably even pattern of neural circuitrary. The neo-cortex initiates and implements a customized set of very common algorithms to completediversified functions of intelligence, very intelligently[46].

HTM is closely related to a type of neural network. However, HTMs working principles at deep level of semantics, are much differ to the term neural network, because it has been applied to a large variety of diversified systems. In HTMs model typicalunit neurons (called cells in HTM), are arranged in the similar way as of neo-cortex, in columns, in layers, in regions, maintained a hierarchy[12, 54].

HTM systemsare trained number of timeson different types of data. HTM dependsstorage pattern andsequences of SDRs. Information storage and accession is logically different from the conventional models, as it is stored with contextual references.Data organization in typical computer memory has a flat and static, without having anyintegratedconcept of time as constraint. A programmer is always free to selectdata structure of his own choice[15] and have full control over stored information answering how and where it is stored. But on the contrary, HTM memory model implements many restrictions. HTM memory has a time-basedhierarchical structure and information is always stored in a distributed manner. Size of hierarchy and basis of system training is always specified by HTM user, but HTM has the control over,“where and how” information is stored[55].Though classic computing is much different from internal structure of HTM networks, but general-purpose computers are used to model these HTMs by incorporating the unique functions of cell hierarchy, cell activation time and sparse distributed representations of cells. Creation of dedicated hardware is also possible for purpose-built HTM networks[4].

Human vision, touch, hearing, language, and behavior lead to inventiveness of HTM properties and principles. Spontaneous and easy to grasp examples of human behaviour are very useful in this regard [14]. More importantly HTM possesses general capabilities. These predictive auto learning machines could be subjected and exposed to number of problems of similar nature. They can also be applied and exposed to purely informational input streams.

Therefore, non-human streams of data could be supplied as their input, such as continuous signals received from radar and infrared. HTMs performance is outclass for web traffic streams, weather data streams and texts from social media streams etc.,[56, 57].

As the derived HTM model is based on the working principles of Neo-Cortex. The basics of this memory are:

1. This is the hierarchy of identical region[58]. This is the memory of sequences, identical patterns or the memory of sequences of sequences as shown in fig 2.4.The uniqueprimarypurpose of all neo-cortical tissues is learning and rememberingorders of patterns [21].

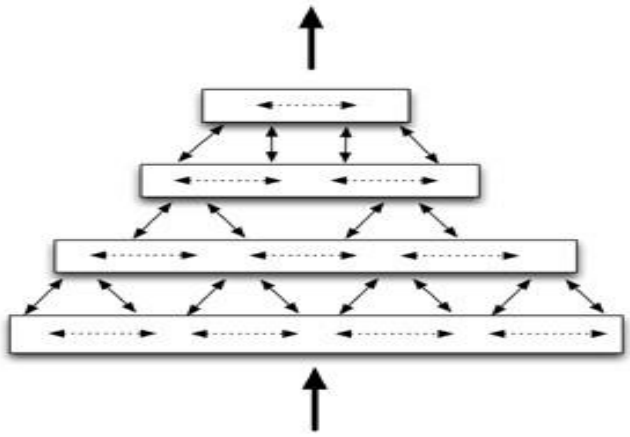

Figure2.4: HTM regions communicating information within levels, (http://www.numenta.com)

2. Each of region learns a sequence[59]. HTM regions are comprised of many cells[60] as shown in fig 2.5.

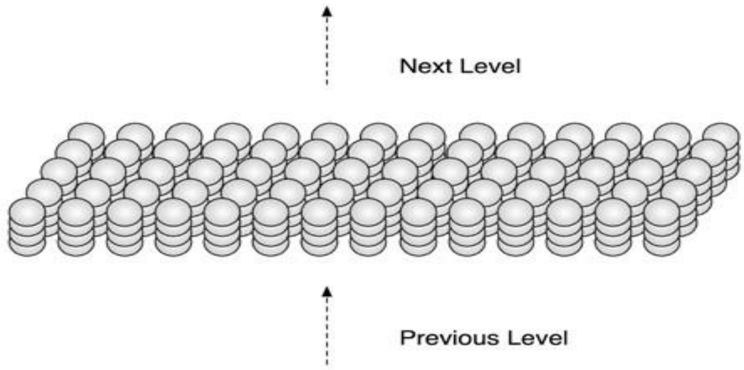

Fig2.5: A segment of an HTM region. Each region is comprised of numerous cells (http://www.numenta.com)

3. Stability increases going up hierarchy in layers, if input is predictable.[8]Activated cells in a HTM region[58] are shown in fig 2.6.

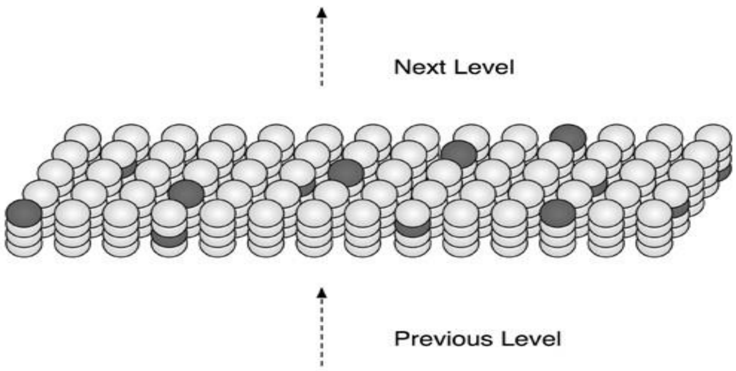

Fig 2.6:SDR of activated cells in an HTM region(http://www.numenta.com)

4. Sequence unfolds going on.[4, 14].

2.5.1Cellular Layers:

The neo-cortex is divided into six layers with scores of different cell types. All type of cells are located in each layer. Layer 6 contains gird cells. Layer 5 contains both cortical neurons and displacement cells used in movement detection as Layer 5 contains thick-tufted neurons [61]. The six layers as shown in fig 2.7, layer 2 and layer 3 form are sequence memory interface. Layer 5 and Layer 6 are feed backward layers while layer4 and layer2/3 are feed forward layers. Layer3 is again subdivided into “a” and “b” depending on the functionality. Each layer of cells is implementing a type of sequence memory. Two forward layers are making inference or pattern recognition[59]. Layer5 have the cells that generate motor behavior and is further sub divided into L5a and L5b. Speech is generated by cells in layer5[62].Layer 6 hasto develop potential in hierarchy and is also divided into L6a and L6b according to their biological structure[5, 10].

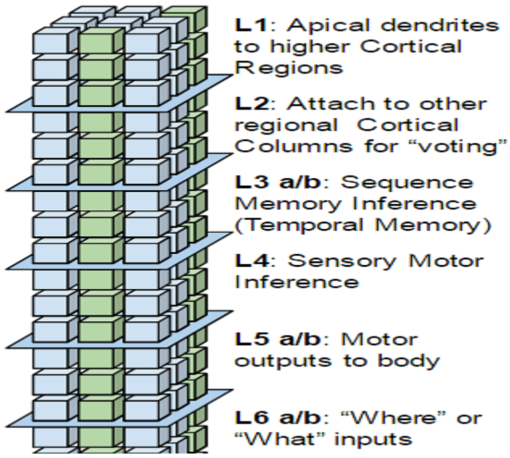

Fig2.7: Layers in neo-cortex and their functions (http://www.numenta.com)

Cortex is not just receiving information from senses but also from motor commands or other behaviours or changes done in the body by other parts of the body. In Layer4 sensory motor inference also takes place. If the input in layer4 is stable, the prediction to layer3 which is high order sequence memory, is stable. These are **universal inference steps** and are applied to all sensory modalities. This Produces receptive field properties seen in cortex[63].

Table 2.1 describes the data type, capabilities and applications of the layers architecture as described in HTM.

Table 2.1: Layer Structure of Cortex showing capabilities and applications of each layer

Layer	HTM Sequence	Data	Capabilities	Applications
2/3	High order Interface	Streaming	Prediction**, Anomaly Detection,** Classification	Information Technology, Security, NLP
4	Sensory Motor Interface	Static with simple behavior	Classification, Prediction	Vision, Image Classification, Network Classification
5	Motor Sequence	Static and Streaming	Goal Oriented Behaviour	Robotics, Smartbots, Proactive Defense
6	Attention, Feedback	Multi Sensory Modalities	Yet to be discovered	Multi Sensory Modalities Multi behavior Modalities.

2.5.2The Neurons:

In biological neuron 10% of synapses are closed to the cell body[14] where feed forward patterns come to and generates spikes on the cell. The other two regions where bosal dendrites and epical dendrites on the top make local connections as shown in fig 2.8. These synaptic connections when repeatedly grown then they make boutons which make permanent memory.

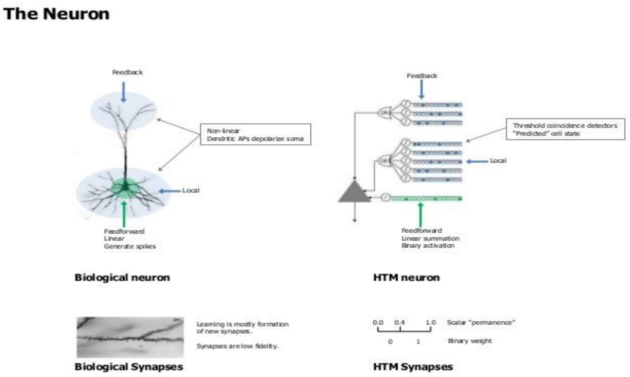

Fig2.8: Comparison of biological neuron with HTM neuron(http://www.numenta.com)

Distance and Direction are encoded by many types of neurons in spatial navigation and memory. This encoded distance and direction is also in the form of binary vector sparsely filled. This information is the base for spatial cognition. The output signals of these neurons are the base of hierarchy for the bottom-up picture of situation andtop-down memory-guided initiation of visuospatial descriptions and directional and route arrangement[64].

2.6Sparse Distributed Representations (The language of intelligence):

Neurons are interconnected multidimensionally in neo-cortex. At certain instant only a limitedfraction of the neurons remains active. Any information is symbolized by these active limited fractions of neuronsinsidelargenumber of neurons. Such coding is termed as a “sparse distributed representation”. The word “Sparse” explains that only a limitedfraction of neurons remains active at certain instant of time. The next word “Distributed” means that the activations of limited number of neurons with in the total relevant neurons is required to represent something[23, 24, 65, 66]

Contextual population of neurons is a ground to understand the meaning conveyed by single active neuron. The input to HTM regions is sparse distributed representations. SDRs is the only internal language of computation within an HTM region’s memory mechanism. HTM region always receive input in the form of a distributed representation,sometimes the input is not sparse.So HTM region converts this input into a sparse distributed representation.Learning, inference and prediction are the functions every HTM region always have to perform [14, 67, 68]

2.6.1 Learning:

Any HTM region learns from inputs the patterns and finds the sequence of this learned pattern. The region is unaware about the input’s representation (meanings). It looks for combinations of input bits, which are called spatial patterns[8]. Furthermore, it tries to determine howthe streams of input bits appear in orderwith passage of time, which are called temporal patterns or temporal sequences[4, 69, 70].

A single HTM region’slearning capability is not perfect. HTM region corrects its learning automatically depending on memory allocated to that region and density of inputs received. On the reduction of allocated memory to a region spatial patternslearned willbe simpler and if the allocated memory is increased spatial patterns learned would be more complex. On arrival of each new input the algorithms developedfor a HTM region,have the capability of continuouslyconnected learning, thus, by not separating inference phase from learning phase despite the fact that inference improves by additional learning. The HTM region will gradually change as the patterns in the input change[3, 23, 71].

HTM continues to learn after initial training. After training phase learning is disabled. The other way tocontinue to learning at the higher levels but not at the lowest levels of the hierarchy is whenat first HTM has learned the fundamentalconstruction of a world, maximum new learning takes place in higher levels of the hierarchy.If the a new input is provided to an HTM or an unseen at low-level structure, HTM will take longer time for learning these new patterns of input streams [3].

2.6.2 Inference:

HTM perform inference on novel inputs after when it has learned any pattern of inputs in. When HTM receives input, it matches it to previously learned patterns. Matching similarityof new inputs to formerly stored sequences is the degreeof inference and pattern matching. Input to HTM regions are SDRs. The use of sequence and inference in a region are more complex. During inference and training an HTM region has the capability of handling new input. Use of sparse distributed representation is away an HTM region copes with new input [25]. An important property of sparse distributed representations is matching a limited portion of input pattern with the confident that the match is weighty[12, 27].

2.6.3 Prediction:

Sequences of ordered patterns are stored in every region of HTM. By matching newly arrived input with stored sequences, a prediction is propagated by HTM, about succeeding inputs [72].Only changes or transitions are stored in the form of sparse distributed representation in an HTM region. These changes or transitionsare like a linear sequence in some cases, but generallynumber of imaginableupcoming inputs are predicted simultaneously. HTM region mademany predictions depending on contextual input mightdrag back far in time. A large portion of HTM memory is dedicated to sequence memory or storing changes of spatial patterns.

2.6.4 Behaviour:

Human behavior influences what he perceives. Almost all his actions change what he senses. Motor behavior and sensory input are intimately entwined[73-75]. Primary motor region in neo-cortex is a single region, where all types of motor commands are initiated in the neo-cortex. It was also revealed with the passage of time that mostof the regions in the neo-cortex have a motor output. Again, it was observed and discoveredthat all cortical regions are capable of receiving inputs of both typesi.e. sensory and motor functions. It isplanned in HTM model that a motor output integratedin regions as a new prediction.

On the basis of HTM, sparse distributed representations (SDRs) or semantic folding technique for data-encoding mechanism[11], are used with the following properties:

- Many bits are used to represent a data item, may be in thousands. Each bit means a neuron. At a certain point active represented by 1, while inactive neurons are represented by 0's.
- Few of them are 1's most are 0's. For example, with in 2000 bits only 2% are active. Sparsity means that most of the neurons are inactive thus represented by 0's.
- Each bit has some meaning known as semantic meaning. Each bit represents a specific feature.
- Meaning are Learned in this representation. Commonly top forty attributes are taken to represent data.

2.6.5 SDR Encoded Data and its Properties:

i. By storing the indices of only "on" bits SDRs can be proficiently stored by minimizing the information loss even if subsampled[13].

ii. Encoding sensor ensures the semantic meaning of each bit in an SDRwithin the context.

iii. SDR encoder ensure that Similar things must look similar, Similarity iscomputed using simple distance measures.

iv. Fault tolerance is the main advantage of SDRs since semantics of any item are preservedby changing or shifting set bits.

v. The resultant union SDR of all other SDRsis the SDRthat comprises all the information of the constituents. For the relevancy determination of unseen upcoming SDR a simple comparison is enough.

vi. "locality-based weighting structure"is used to determine the sparsity of an SDRs in a semantically reliable fashion.

2.7Use of SDR's for Sequence Learning

The fundamental difference between HTM sequence memory and former sequence learning models is the use of sparse distributed representation[65, 66, 76]. Information representation is done by activating of a certain number of neurons in certain instant time in cortex is called as sparse coding[1, 16, 17, 19, 77]. Temporal sequences with the help of SDRs are represented in HTM sequence memory. According to mathematical properties of SDRs [8, 11, 14], each and every neuron in the HTM modelcapable to learn as well ascategorize a large set of patterns in noisy and uncertainenvironments [13]. In HTM sequence memory arobust distributed neural representation called binary SDRs for temporal sequences comes out from computation. This sparse representation is value able for many tasks like anomaly detection and classification. [78].

Flexible coding system is significantlyvital, for online streaming analytics, due to the numeralunique symbols which are unknown before. In such cases, the coding scheme range requires change without affecting previous learning of the model ultimately leading to algorithm that uses a flexible coding systemcapable to denote a large number of uniquesymbols or aextremely extended data range [14, 23, 76]. The SDRs used in HTM allow number of representations fordifferent predictions with minimum of collision chances. The succeeding generation of neural networksmust have these propertiesof SDRsfor an ideal coding format[12].

2.8 Encoding:

Data received from various sources can be of various types and formats like: real numbers, integers, categories; numbers, date, time, color, weight, height ... Commonly the format received from any input source is supplied to Neural Network world with or without little bit reshaping or changing format. we can sometimes use the data as it is or slightly reshape it. In case of HTM the data format is completely changed by converting data to SDRwhichcannot be doneautomatically. Semantics or hidden relations between data items are extracted and explored to encode them into SDR with their specific position.Each bit in an SDR have its own semantic meaning. A concept is represented with the combination of bits having some spatial or temporal relation. This is a manual process. There is no general technique of extracting semantic information from any data source. Once semantics are extracted, the encoders are built to convert data into SDR format.

2.9 Semantic Folding Theory:

Conversion of common language words using their semantic meanings into a sparse binary vector is called semantic folding. A conversion method of any language from its text representation into aclear and semanticalrepresentation that can be processed by HTM networksis called semantic folding.

Sparse binary representation solves complex NLP problems[33, 65, 76, 78-80]. Simple Boolean operators are used in processing. Similarity function like Euclidian Distance are used. Semantic folding provides solution to many statistical NLP problems. Precision and recall etc., can easily be solved by applying Semantic Folding.

2.9.1 Origins and Goals of Semantic Folding Theory

SFT provide a framework to unfold natural language perception and productionusing semantic information operated and manipulatedin the neo-cortex, during initial language acquisition phases.Data representational problem was solved by applying a unique approach of word semantics. This approach is based upon words meaning representation in way to be processed and computed by cortical learning algorithm. Language processing is done with the help of its meaning as humans do.This enables a high performance in of natural language processing.

A highly technical memory structure for cortical processing and cortical memory is the Content Addressable Memory (CAM). Standard memory and CAMare integrated and linkage with memory cells established with address input and data input/output. The limitation of CAM is that it is content addressable memory. The data is converted as address and moved as address input.In real time processing the amount of available memory modules is directly proportional to the volume of data, is the second big advantage of the HTM-CAM.

Data storage in memory cells is dependent of the mechanism of cortical memory to generate addresses.At each time of data storage, memory cell address is captured. Memory cell implementation is minimal provided by corticalmemory function, with semantic data storage. The result of any “query” isfound immediately having extended width.

2.10 Semantic Fingerprinting:

In machine learning area, the conceptual foundations of Semantic Folding Theory and HTM Theory are same by aiming toimplement the latest findings of neuroscience. Both systems work in such a way that SemanticFolding acts as encoder for the incoming data stream and HTM’s implementation performs as intelligent backend. The task of text data conversion into Sparse Distributed Representation (SDR) done through Semantic Folding. It also enables the similarity operation and Boolean operations on these SDRs. This data is sent to an HTM implemented model, as Sparse Distributed Representations[23]. As discussed earlier, SDRs are binary vectors, sparsely filled (with very few bits turned on), every bit representing specific semanticevidence (i.e.specific meaning of that corresponding bit).HTM theory explains that the human neo-cortex is a memory system and not a processor, for SDR pattern sequences. When an input data stream enters to HTM layer, it generates next prediction of next incoming SDR pattern[12, 27, 78] based on input it received. In start, the predictionsdiffer from the actual data but afterwards HTM layer converges to make relevant and correct predictions. This capability of prediction explainsnumber of behavioral indicatorsfor intelligence[56, 71, 78].

2.11Semantic Folding:

For the implementation of HTM to a problem, essentially, given input data is converted into the SDR format with the following Characteristics.

- A very huge binary vectors are termed as SDR with millions of bits.
- The total number of bits with status “on” in an SDR is very limited at a specific time interval.
- In SDR format similar data always resembles.

- In any SDR every bit represents a specific meaning with respect to its context.
- The resultant SDR which is union of many SDRs possesses complete information of its component SDRs.

Semantic Folding process consists of the following steps:

- Semantic universe for system is the collection of definition of reference text corpus (reference terms) of documents in which system is supposed to work [23]. SF system have the knowledge of vocabulary and use it as in Language Definition Corpus (LDC).
- A term snippet is used for cutting down the text document of LDC into small text snippets having a single context.
- A 2D semantic map is created by referred collection of snippets[23], in such a way that snippets with irrelevant topics are placed at a distantfrom each on the map and snippets of same context are placed near each other.
- Creating a corpus by word list Creation is the next step.
- Semantic fingerprint is generated with the help of this list from top to bottom as;
  All context of a word are set to 1.
  A large sparsely filled binary vector is prepared for each word.
  This vector is called semantic fingerprint.
  The structure of 2D map is folded into each word representation.

2.11.1 The Database:

In Semantic Folding the Database is collection of words. There are, text snippets ranging from one to several on matrix. The integral words located in the semantic space represent topic. This is based on the analogy that the neo-cortex is arranged as a two-dimensional grid.

2.11.2 Language Definition Corpus:

For representation in language definition corpus if documents are selected from a specific domain, the said domain specific database would be the result. If, documents from the PubMed archive are selected the resulting database will cover medical English. Collection from twitter will make a LDC of “twitterish” database. Similarly, is the case of other languages. Bias associativity in a database depends upon the size of the generated text snippets. If the snippets are small, (1-3 sentences), the associativity will be fewer. The larger the text snippets lead to number of associated concepts. Practically, problem domain decides about bias level. Proposed system uses the yelp dataset as database and is labelled as ‘yelp-database’.

2.11.3 Definition of a General Semantic Space:

A database for every desired language would be generated in order to attain cross language functionality, keeping forth the primary semantic space topology, the generated fingerprint for a single perceptedwould be the same in all languages. For a given domain semantic space tuning is done by generating fingerprints and using available area, by improving the semantic resolution. The word Tuning stands for selecting relevant representative training material. Domain expert is responsible for this task.

2.12 Combination of database with HTM Model (HTMLA):

Earlier, before the development of HTM theory, computer scientists dreamed that human language be accessible to computer programs. Human expertise in overcoming the ambiguities, irregularities of natural languages failed all the statistics or probability calculus, and other mathematical approaches. To overcome these problems following experiment (as shown in fig: 2.9) was carried out:

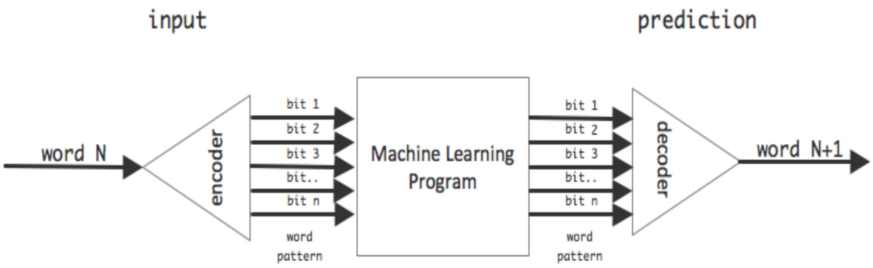

Fig.2.9:experimental setup overview[81]

- With a number of binary inputss amachine learning (ML) program has been taken. This ML program has been trained on binary sequence of patterns in a certain time. The ML program gives outputs in a manner of prediction to antedatenext pattern,
- A program named Encoder,converts an English word into a ML readable binary pattern and reconverts any ML readable binary pattern into the English word. The main quality of the Encoder is that it convertsthe words into similar binary patterns which have the same semantics and vice versa. Euclidian distance is the measure used to find similarity ratio between two binary patterns.
- The encoder converts and encodes every word into a binary vector of 16kbits pattern.
- The configure machine learning program allows input as well as output prediction vectors of length 16K bits patterns.
- The encoder with the help of machine learning program has formed a system that takes input in form of words and returns output in the form of words as well.
- The encoder that converts "word to pattern" linked to the machine learning program inputs sequences of words for which machine learning program outputs a "binary pattern" representing a"prediction" of next expectation.
- The decoder that converts pattern to word linked to the prediction outputs of the machine learning program.
- So, a sequence of natural language words is supplied to the system generating predictionsof next probable word as its result based on all precedinginput series.

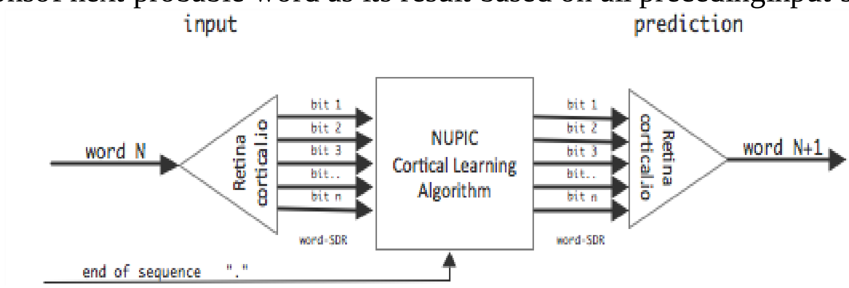

Fig. 2.10:experiment implementation[81]

- Hierarchical Temporal Memory (HTM) is the machine learning program used in this experiment and this was the first model developed by Numenta in 2011 as fig 2.10 shows. This experiment motivated us for this research.
- It is clear that initially an uncertain programis trained on series of input data patterns and is capable to predict a resultas pattern created on the basis of previous input of sequences of patterns.
- In above experiment data of English language is used. English wordswere encoded into binary patterns and were directly brought as an input into an online learning algorithm, the HTM LA (learning algorithm) learns from input data.
- The learning algorithm generates a binary predicted vector of size as the input vector, decoded into a word.
- The full stop is understoodas an end signal by the HTM LA. It also informs the system that a new series of patterns to be started for next sentence.

3 Advantages:

Advantages of HTM model and Semantic Folding theory are as under:

1. To operate this model no special assistances are required.
2. This system works on unsupervised learning rules.
3. This system works on data driven approach, only domain experts are sufficient to train the system.
4. Semantic Folding features allow a best representation of concepts.
5. System learns semantic features from the language model itself which reduces semantic bias.
6. The expressive structures are clear and have sufficient semantic base making the system reliable by interpreting the generated results.
7. Drastic reduction of vocabulary disproportion, false positive results are minimized.
8. “Binary encoded semantics form” improve the processing speed over old-style methods.
9. Processing pipeline implementation is easy due to having all semantics of same size.
10. Query response time minimizes having all pre-calculated semantic representations.
11. Independent calculations are used by algorithms.
12. Performance improvements are achieved by implementing similarity algorithms in hardware reducing search time.
13. By using the associated semantic space for different languages, the resulting fingerprints become language independent. This means that fingerprint of any language could be matched with any other language. The filter criterion considered for English could be used with all other languages.

4 Summary:

The HTM cortical learning algorithm forms a fundamentalstructure of neural organization in the neo-cortex. This explains the multi-dimensional connectivity of neurons and their learning sequences of sparse distributed representations. Learning changes in the different layers of neo-cortex are utilized for related but different purposes. Different regions of neo-cortex have feed-forward inputs.With the combination of layer 3 and layer 4 cells neo-cortex creates a sparse distributed representation of the input. Cells in layers 2, 3, 4, 5, and 6 share this sparse distributed representation. This is done enforcing all cells of the columns of down layers to respond to the same feed-forward input. First order temporal sequence is learned in Layer 4 cells, when they are present. They use the HTM cortical learning algorithm which create representations that are constant or invariant to spatial sequences.Layer 3 cells use the HTM cortical learning algorithm to learn variable-order temporal changes and form stable representations that are passed up the cortical hierarchy. Layer 5 cells learn changes with timing. However, layer 2 and layer 6, due to the horizontal connectivitylearn some form of sequence memory.

Sparsity of active neurons lead to the development of SDR which is the basic learning and operating language of HTM based semantic folding theory that gives the solution to functions of layers of neo-cortex as identified by Hawken's model.

Conclusion:

Upon the working principles of human neo-cortex, the Hierarchical Temporal Memory (HTM) model has been developed by Jeff Hawkins, which is a proposed theoretical framework for sequence learning. Both types of data numerical and categorical are best suited input types for HTM model'sworking. Semantic Folding Theory (SFT) is based on HTM to represent a data stream for processing in the form of Sparse Distributed Representation (SDR). SFT offers a framework for unfolding how semantic information is manipulated for natural language observation and creation, towards the details of semantic foundations during initial language learning phase.

References

1. Rozado, D., F.B. Rodriguez, and P. Varona, *Extending the bioinspired hierarchical temporal memory paradigm for sign language recognition.* Neurocomputing, 2012. **79**: p. 75-86.
2. Padilla, D.E., R. Brinkworth, and M.D. McDonnell. *Performance of a hierarchical temporal memory network in noisy sequence learning*. in *2013 IEEE International Conference on Computational Intelligence and Cybernetics (CYBERNETICSCOM)*. 2013. IEEE.
3. Hawkins, J.C., et al., *Temporal memory using sparse distributed representation.* 2019, Google Patents.
4. Hawkins, J., S. Ahmad, and Y. Cui, *Why does the neocortex have layers and columns, a theory of learning the 3d structure of the world.* bioRxiv, 2017: p. 162263.
5. Yu, C.-H., Z. Qin, and M.J. Buehler, *Artificial intelligence design algorithm for nanocomposites optimized for shear crack resistance.* Nano Futures, 2019. **3**(3): p. 035001.
6. Bard, N., et al., *The Hanabi challenge: A new frontier for AI research.* Artificial Intelligence, 2020. **280**: p. 103216.
7. Mauk, M.D. and D.V. Buonomano, *The neural basis of temporal processing.* Annu. Rev. Neurosci., 2004. **27**: p. 307-340.
8. Clegg, B.A., G.J. Digirolamo, and S.W. Keele, *Sequence learning.* Trends Cogn Sci, 1998. **2**(8): p. 275-81.
9. Basseville, M. and I.V. Nikiforov, *Detection of abrupt changes: theory and application*. Vol. 104. 1993: prentice Hall Englewood Cliffs.
10. Hawkins, J., et al., *A framework for intelligence and cortical function based on grid cells in the neocortex.* Frontiers in neural circuits, 2019. **12**: p. 121.
11. Kanerva, P., *Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors.* Cognitive computation, 2009. **1**(2): p. 139-159.
12. Cui, Y., S. Ahmad, and J. Hawkins, *Continuous online sequence learning with an unsupervised neural network model.* Neural computation, 2016. **28**(11): p. 2474-2504.
13. Cui, Y., S. Ahmad, and J. Hawkins, *The HTM spatial pooler—A neocortical algorithm for online sparse distributed coding.* Frontiers in computational neuroscience, 2017. **11**: p. 111.
14. Hawkins, J. and S. Ahmad, *Why neurons have thousands of synapses, a theory of sequence memory in neocortex.* Frontiers in neural circuits, 2016. **10**: p. 23.

15. Hawkins, J.C., et al., *Architecture of a hierarchical temporal memory based system*. 2013, Google Patents.
16. Amunts, K., et al., *The Human Brain Project—Synergy between neuroscience, computing, informatics, and brain-inspired technologies.* PLOS Biology, 2019. **17**(7): p. e3000344.
17. Kasinski, A.J. and F. Ponulak. *Experimental Demonstration of Learning Properties of a New Supervised Learning Method for the Spiking Neural Networks*. in *ICANN*. 2005.
18. Martin, K.A. and S. Schröder, *Functional heterogeneity in neighboring neurons of cat primary visual cortex in response to both artificial and natural stimuli.* Journal of Neuroscience, 2013. **33**(17): p. 7325-7344.
19. Yen, S.-C., J. Baker, and C.M. Gray, *Heterogeneity in the responses of adjacent neurons to natural stimuli in cat striate cortex.* Journal of neurophysiology, 2007. **97**(2): p. 1326-1341.
20. Brea, J., W. Senn, and J.-P. Pfister, *Matching recall and storage in sequence learning with spiking neural networks.* Journal of neuroscience, 2013. **33**(23): p. 9565-9575.
21. Hawkins, J. and S. Blakeslee, *On intelligence: How a new understanding of the brain will lead to the creation of truly intelligent machines*. 2007: Macmillan.
22. Yoshimura, Y., et al., *Properties of horizontal and vertical inputs to pyramidal cells in the superficial layers of the cat visual cortex.* Journal of Neuroscience, 2000. **20**(5): p. 1931-1940.
23. Webber, F.D.S., *Semantic folding theory and its application in semantic fingerprinting.* arXiv preprint arXiv:1511.08855, 2015.
24. Földiák, P., *Sparse and explicit neural coding.* Principles of neural coding, 2013: p. 379-389.
25. Hawkins, J., S. Ahmad, and D. Dubinsky, *Cortical learning algorithm and hierarchical temporal memory.* Numenta Whitepaper, 2011. **1**: p. 68.
26. Ahmad, S. and S. Purdy, *Real-Time Anomaly Detection for Streaming Analytics.* 2016.
27. Cui, Y., et al. *A comparative study of HTM and other neural network models for online sequence learning with streaming data*. in *2016 International Joint Conference on Neural Networks (IJCNN)*. 2016. IEEE.
28. Hawkins, J.C., I.R. Marianetti, and A. Raj, *Encoding of data for processing in a spatial and temporal memory system*. 2014, Google Patents.
29. Numenta, *numenta.org.* Numenta Whitepaper, 2014: p. 68.
30. Park, W.J. and J.-B. Park, *History and application of artificial neural networks in dentistry.* European journal of dentistry, 2018. **12**(4): p. 594.
31. Deng, L., G. Hinton, and B. Kingsbury. *New types of deep neural network learning for speech recognition and related applications: An overview*. in *2013 IEEE international conference on acoustics, speech and signal processing*. 2013. IEEE.
32. Suk, H.-I., *An introduction to neural networks and deep learning,* in *Deep Learning for Medical Image Analysis*. 2017, Elsevier. p. 3-24.
33. Pathirage, C.S.N., et al., *Structural damage identification based on autoencoder neural networks and deep learning.* Engineering structures, 2018. **172**: p. 13-28.
34. Mendez, K.M., D.I. Broadhurst, and S.N. Reinke, *The application of artificial neural networks in metabolomics: A historical perspective.* Metabolomics, 2019. **15**(11): p. 142.
35. Mitchell, T., et al., *Never-ending learning.* Communications of the ACM, 2018. **61**(5): p. 103-115.
36. Dietterich, T.G., *Machine-learning research.* AI magazine, 1997. **18**(4): p. 97-97.
37. Williams, C.K. and D. Barber, *Bayesian classification with Gaussian processes.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998. **20**(12): p. 1342-1351.
38. Rasmussen, C.E. *Gaussian processes in machine learning*. in *Summer School on Machine Learning*. 2003. Springer.
39. Schnelle, H., *Language in the Brain.* 2010: Cambridge University Press.
40. Vinje, W.E. and J.L. Gallant, *Natural Stimulation of the Nonclassical Receptive Field Increases Information Transmission Efficiency in V1.* The Journal of Neuroscience, 2002. **22**(7): p. 2904-2915.
41. Pietron, M., M. Wielgosz, and K. Wiatr. *Parallel Implementation of Spatial Pooler in Hierarchical Temporal Memory*. in *ICAART (2)*. 2016.

42. Polsky, C., *Putting space and time in Ricardian climate change impact studies: agriculture in the US Great Plains, 1969–1992.* Annals of the Association of American Geographers, 2004. **94**(3): p. 549-564.
43. Poirazi, P., T. Brannon, and B.W. Mel, *Pyramidal neuron as two-layer neural network.* Neuron, 2003. **37**(6): p. 989-999.
44. Larkum, M.E. and T. Nevian, *Synaptic clustering by dendritic signalling mechanisms.* Current opinion in neurobiology, 2008. **18**(3): p. 321-331.
45. Choi, W., et al., *Tomographic phase microscopy.* Nature methods, 2007. **4**(9): p. 717-719.
46. Major, G., M.E. Larkum, and J. Schiller, *Active properties of neocortical pyramidal neuron dendrites.* Annual review of neuroscience, 2013. **36**: p. 1-24.
47. Schiller, J. and Y. Schiller, *NMDA receptor-mediated dendritic spikes and coincident signal amplification.* Current opinion in neurobiology, 2001. **11**(3): p. 343-348.
48. Schiller, J., et al., *NMDA spikes in basal dendrites of cortical pyramidal neurons.* Nature, 2000. **404**(6775): p. 285-289.
49. Moser, E.I., E. Kropff, and M.-B. Moser, *Place Cells, Grid Cells, and the Brain's Spatial Representation System.* Annual Review of Neuroscience, 2008. **31**(1): p. 69-89.
50. Doeller, C.F., C. Barry, and N. Burgess, *Evidence for grid cells in a human memory network.* Nature, 2010. **463**(7281): p. 657-661.
51. Domnisoru, C., A.A. Kinkhabwala, and D.W. Tank, *Membrane potential dynamics of grid cells.* Nature, 2013. **495**(7440): p. 199-204.
52. Stensola, H., et al., *The entorhinal grid map is discretized.* Nature, 2012. **492**(7427): p. 72-78.
53. Lewis, M., et al., *Locations in the neocortex: A theory of sensorimotor object recognition using cortical grid cells.* Frontiers in Neural Circuits, 2019. **13**: p. 22.
54. Kedra, J., et al., *Current status of use of big data and artificial intelligence in RMDs: a systematic literature review informing EULAR recommendations.* RMD open, 2019. **5**(2): p. e001004.
55. Spruston, N., G. Stuart, and M. Häusser, *Principles of dendritic integration.* Dendrites, 2016. **351**(597): p. 1.
56. Rah, J.-C., et al., *Thalamocortical input onto layer 5 pyramidal neurons measured using quantitative large-scale array tomography.* Frontiers in neural circuits, 2013. **7**: p. 177.
57. Bühler, K., *Theory of language.* The representational function of language, 1990.
58. Behrens, T.E., et al., *What is a cognitive map? Organizing knowledge for flexible behavior.* Neuron, 2018. **100**(2): p. 490-509.
59. Burak, Y. and I.R. Fiete, *Accurate path integration in continuous attractor network models of grid cells.* PLoS Comput Biol, 2009. **5**(2): p. e1000291.
60. Giocomo, L.M., et al., *Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing.* Science, 2007. **315**(5819): p. 1719-1722.
61. Thomson, A.M., *Neocortical layer 6, a review.* Frontiers in neuroanatomy, 2010. **4**: p. 13.
62. Dellanna, L., *The function of Layer 5: a Recognizer of Needs for Manipulations.*
63. Guo, T., et al. *Mimicking natural neural encoding through retinal electrostimulation*. in *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*. 2017. IEEE.
64. Bicanski, A. and N. Burgess, *Neuronal vector coding in spatial cognition.* Nature Reviews Neuroscience, 2020. **21**(9): p. 453-470.
65. Foldiak, P. and D. Endres, *Sparse coding.* 2008.
66. Vinje, W.E. and J.L. Gallant, *Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision.* Science, 2000. **287**(5456): p. 1273-1276.
67. Garnelo, M. and M. Shanahan, *Reconciling deep learning with symbolic artificial intelligence: representing objects and relations.* Current Opinion in Behavioral Sciences, 2019. **29**: p. 17-23.
68. Ahmad, S. and J. Hawkins, *Untangling Sequences: Behavior vs. External Causes.* BioRxiv, 2017: p. 190678.
69. Sejnowski, T.J., *The unreasonable effectiveness of deep learning in artificial intelligence.* Proceedings of the National Academy of Sciences, 2020.

70. Hawkins, J. *What the brain tells us about the future of silicon*. in *2015 IEEE International Electron Devices Meeting (IEDM)*. 2015. IEEE.
71. Marianetti, I.R., A. Raj, and S. Ahmad, *Time Aggregation and Sparse Distributed Representation Encoding for Pattern Detection*. 2014, Google Patents.
72. Clarke, J.M. and E. Zaidel, *Anatomical-behavioral relationships: corpus callosum morphometry and hemispheric specialization.* Behavioural brain research, 1994. **64**(1-2): p. 185-202.
73. Markov, N.T., et al., *Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex.* Journal of Comparative Neurology, 2014. **522**(1): p. 225-259.
74. Raynaud, S., *Cesalli, Laurent/Friedrich, Janette (ed by/Hrsg.) 2014. Anton Marty & Karl Bühler Between Mind and Language-Zwischen Denken und Sprache-Entre pensée et langage. Basel: Schwabe Verlag.* 2016.
75. Transaction, S., *Cho CS, Kooby DA, Schmidt CM, Nakeeb A, Bentrem DJ, Merchant NB, Parikh AA, Martin RC 2nd, Scoggins CR, Ahmad SA, Kim HJ, Hamilton N, Hawkins WG, Weber SM. Ann Surg. 2011; 253 (5): 975-980. In the article "Laparoscopic Versus Open Left Pancreatectomy: Can Preoperative Factors Indicate the Safer Technique?" Table 3 should have presented the data for Analysis of Variables Associated with Major Complications After OLP, and not Analysis of Variables Associated with Major Complications After LLP. The authors regret the error.(See corrected Table 3).* 2011.
76. Cheng, H., *Sparse Representation, Modeling and Learning in Visual Recognition.* Advances in Computer Vision and Pattern Recognition, 2015: p. XIV, 257.
77. Leake, M., et al. *Effect of spatial pooler initialization on column activity in hierarchical temporal memory*. in *Twenty-Ninth AAAI Conference on Artificial Intelligence*. 2015.
78. Hawkins, J.C., et al., *Temporal memory using sparse distributed representation*. 2015, Google Patents.
79. Zhou, S., et al., *Spatial–temporal convolutional neural networks for anomaly detection and localization in crowded scenes.* Signal Processing: Image Communication, 2016. **47**: p. 358-368.
80. Niemasik, J. and D. George, *Supervision based grouping of patterns in hierarchical temporal memory (HTM)*. 2012, Google Patents.
81. Webber, F.E.D.S., *Semantic Folding Theory.*