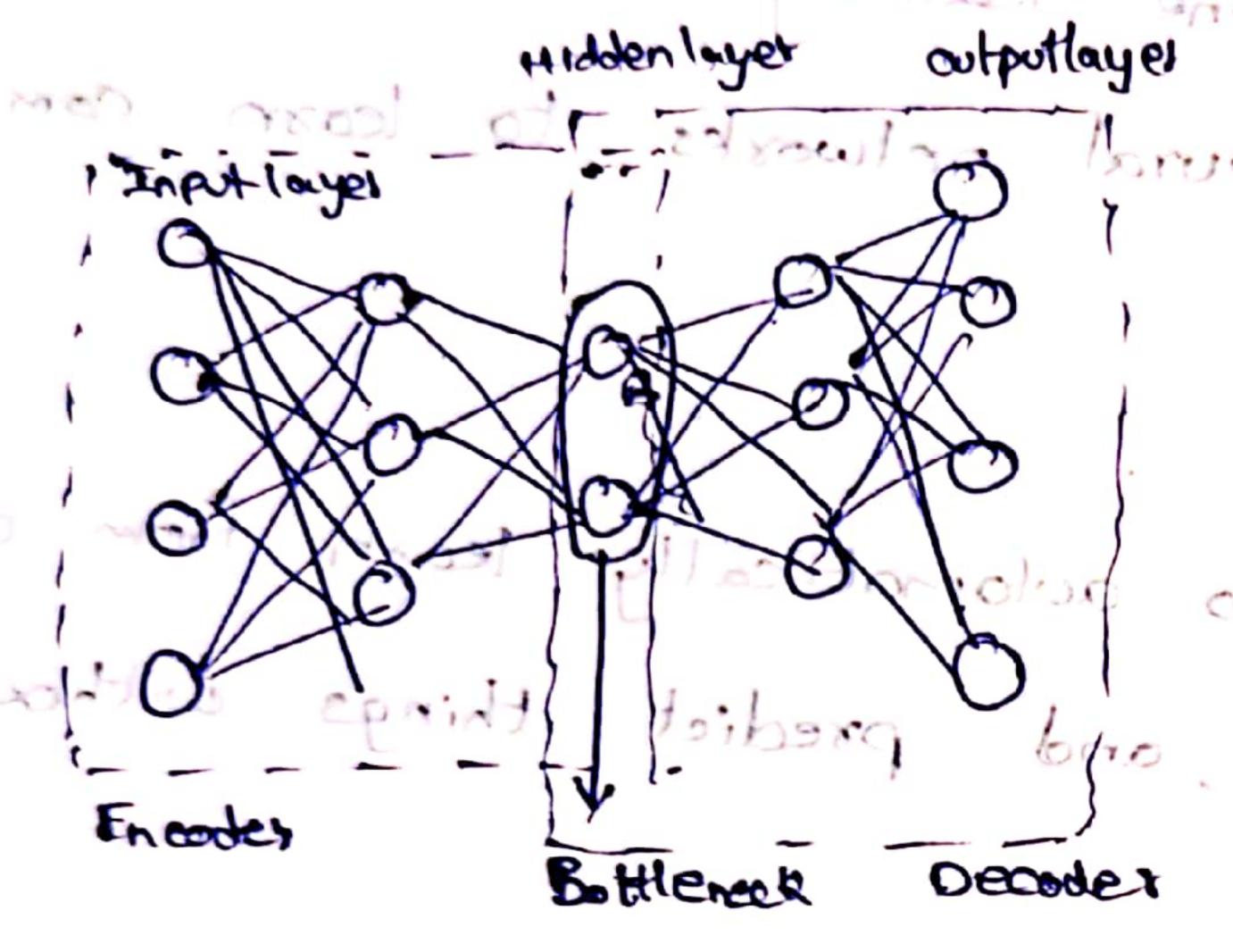
of computer science aims to create intelligent Astificial intelligence: The field machines that on think and function like humans. machine learning! A subset 6x) subfield of artifical intelligence that focus on developing algorithms, and models from data mather than being explicitly Deep learning! At subfield of machine learning that uses multi-Patterne in data. Machine learning enables and machine to automatically learn from data, improve performance from experiences, and predict things without being explicitly programmed. A machine learning system learns from historical data, builds the prediction node's newhonever it receives new data, predicts, the output for it. The accuracy of the predicted output depends upon the amount of data, as the hype amount of data helps to ibuild man have made which predicts the output morre accurately. Invoisionamib - is and a dri add Irput post of Training machine learning Building Logical - Output Models Togal and Models dota strain som data as similar di This layer (thingsly has a lower dimensionality than the infinit and southers are print is not been to the popular

And enclodes is a type of Artificial neurol network that is used for unsupervised learning. The aim of an autoencoder is to learn a lower dimensional representation (enouding) for a higher dimensional de

typically for dimensionality reduction; By straining, the network to apture

the most important posts of the input image.



eg of autoencodes! -) Dimensionality Reduction

Feature extraction

performance trom experiences

- -) Image Dehoising, Compression
- and Don-) Anomaly, Detection princes

Enoder! - The encoder part of the autoencoder is responsible for the compressing the imput doba into lower-dimensional representation compressing the compressing compresses of one or more whidden abuses the standard the most standard to transform the most that use non-linear diactivation, functions to transform the most that use non-linear diactivation, space.

dela into a lower-dimensional space.

Boltleneck: The Bottleneck byer is most pimportant of an autoenceder as it connects the encoder and decoder points of an outcomeder This layer typically has a lower dimensionality than the input and output layers and is used to x represent of the input data.

Decodes in the decoder part of the autoencoder is responsible for the reconstructing the original input data from the compressed representation created by the encoder. The output layer of decodes should have the same dimensionality as the inputilages to set 4 hyperparameters before training an autoencoder: lode size! most imp hyperporameter used to tune the autoencoder. The bottle neck size decides how much the data has to be compressed. "| Number of layers in the indiates depth of encodes & decodor, while a higher be allowed to be a layers in the indiates model complexity, a lower depth is faster to depth increases model complexity, a lower depth increase to Process, and adjillions to slevers was A Mosof nodes por layer: defines the weights we use pier layer. Typically the noof nodes in an encoder decrese till the bottle neck and increase in an decider from the bottleneck. Reconstruction loss: The loss function mercise, to train the autoencoder is highly idependent on the type of input and output time want the autoencoolex adopt to. 4) Denoising autoencodes & of autoencoders: 1) un dércomplete autoencoder 5) Variational autoencoder 2) sparge autoencoder 3) Contractive autoencodes type of autoencoder that is designed to bridional autoencoder (VAE): is a appeared representations. Instead of encoding the input data into a fixed point ha latest space, the VAE encodes the input data into probability distrubtion over the latent space. Used for data compression, synthetic data creation, etc.

A Deep autoencoder is simply an autoencoder with multiple hidden la in both the encoder and decoder. It is capable of learning more Complex representations of input data composed to a shallow autoencode with only one or two hidden layers.

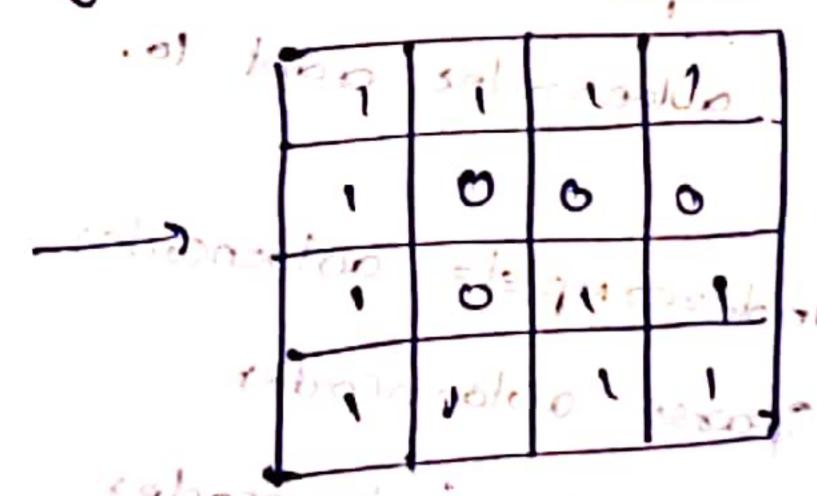
CNN:

CNN stands for convolutional neuralm network, which is a type of A Commonly used for image and video processing tooks such as object operagnition, scene reconstruction, and image classification.

A CUN consists of multiple layers, each designed to postorm a specific operation on when input data. The four import layers in CNN are;

i) Convolution layer: This layer has multiple filters that perform the convolution

operation to extract valuable features from an image. Eve image is considered as matrix of pixel values, whose values is either o(61) 1 Each files slides over the image matrix and perform element were multiplication & sum up the results to produce single value in feature map.

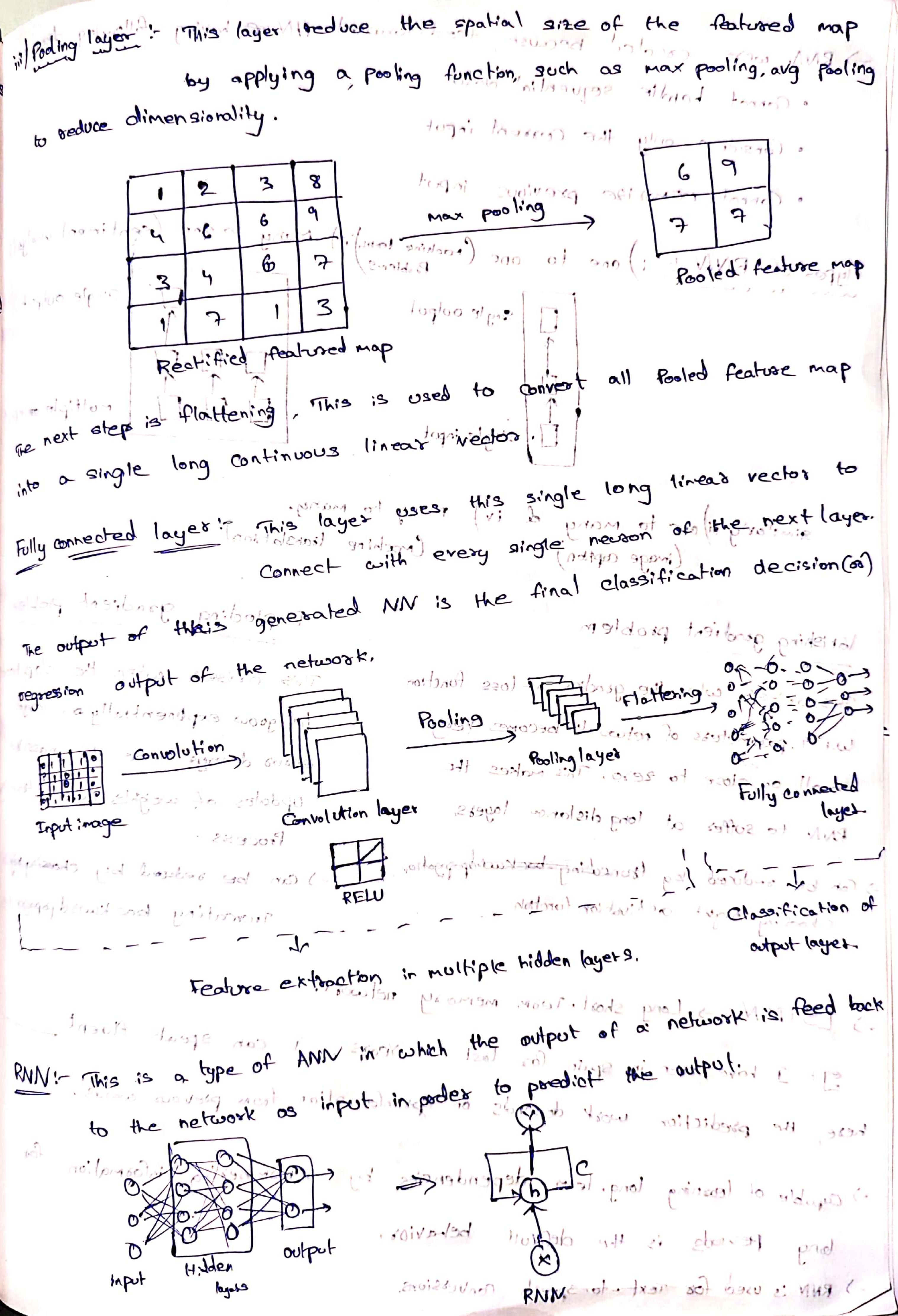


U layeri- This layer applies rectified linear unit activation foreton to the output of the convolution layer. Relu sets all -re values to a &

the values bemains unchanged,

alot sitalt The receipted olds to best - R(2)=max (0,2)

alduriab prillidoduces of in olab togri odi zobenna

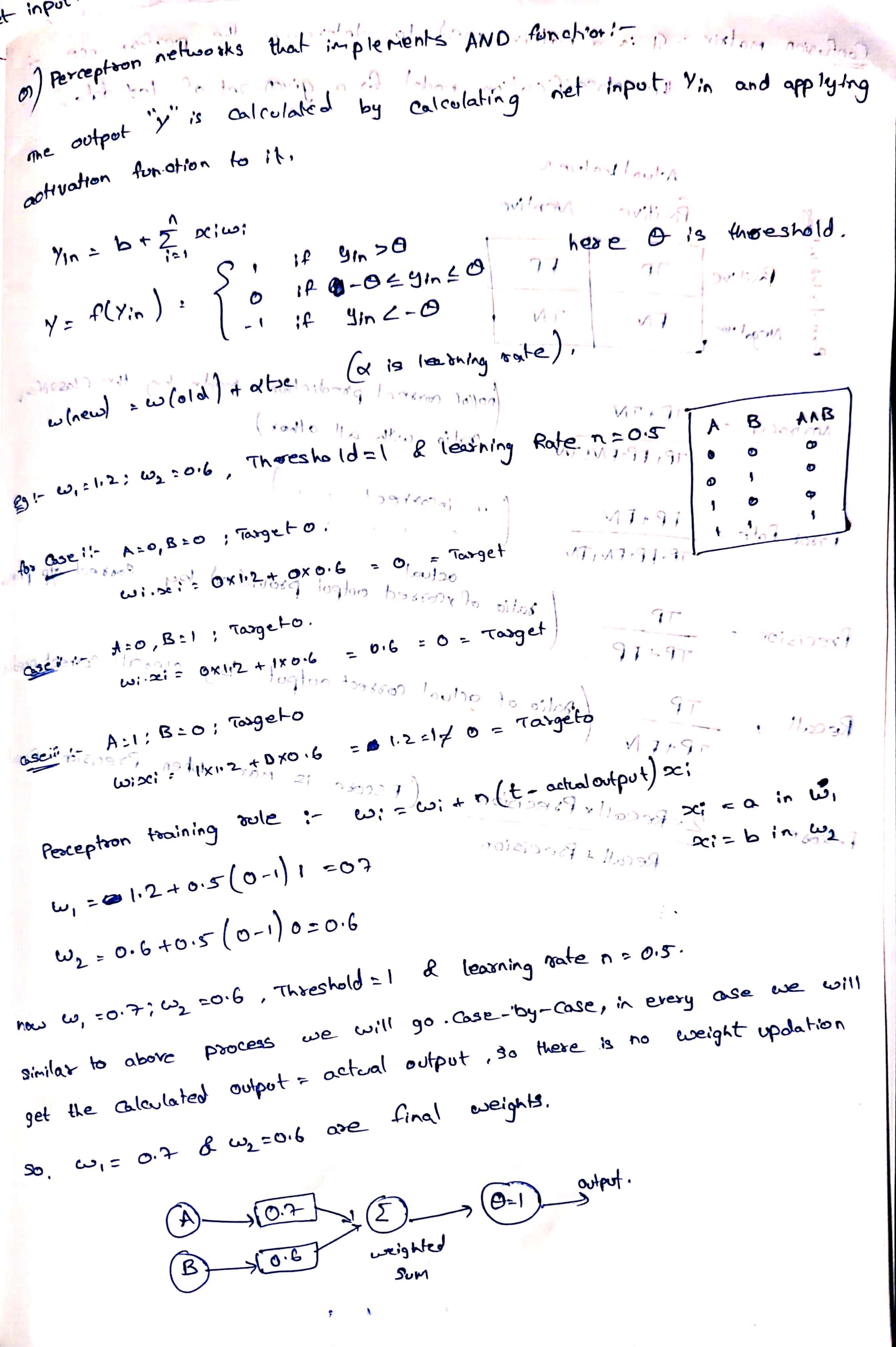


single notpot single input single outpot multiple single single outpot multiple single single outpot single outpo	
Cannot hardle sequential data Canada any the current input Cannot recrotize previous input Types of RNN: James of RNN: Ja	-> RNN were created because, feed-forward NN a having following drawbas
Ganaters only the corrent input Cannot removine previous input Types of RNN ::) and to one (tradine law):) Many to one (centiment analysis) Single output Single input Single output This arises when the gloadiest of loss function to grow explorentially as large sonal or close to zero. This nakes the sorrest decays cause the massive sorrest action sorrest cause the massive sorrest decays sorrest action sorrest cause the massive sorrest decays sorrest cause the massive sorrest cause the	· Cannot hardle sequential data
Cannot renorise previous input Types of RNN:) one to one (realize tous). It have to sine (sentiment analysis single output single output single input single output s	
single input si	· Cannot memorize previous input
single input (machine translation) Exploding gradient problem In gradient problem Exploding gradient problem	Types of RNN: ;) one to one (machine land) is) Many to one (sentiment analysis
similarly ii) one to many d iv) many to many. (image apthon) (image apthon) (machine translation) Exploding gradient problem This arises when the slop to grow explorentally as large carros decays cause the carros decays cause the nousive PRIN to suffer at long distance layets Process: On be reduced by truncating because problem Frocess: On be reduced by truncating backward proposition Frocess: I STMM -> Long short-Term memory network: Exploding store the slop to several carros decays cause the several dependent of the results of the problem of the proble	in a man of the second of the
Similarly ii) one to many d iv) many to many. (image captura) (image cap	Moltable mino.
W.R.T possemeters of network becomes very to grow explorentially as large small or close to zero. This makes the errors decays cause the massive punt to suffer at long distance layers updates of weights during, think is to suffer at long distance layers. San be reduced by truncating beckmand population. Chasting sight archivation foretain Truncating backmand pergating truncating backmand pergating and the second pergating to the property. Long short-Term memory network: egt I have been in spain for last 10 years. I can speak Awent here, the prediction work depends on sprinformation from previous words. Apoble of learning long-term dependencies by remembering information for lang Perreds is the default behavior.	single input
W.R.T possemeters of network becomes very to grow explorentially as large small or close to zero. This makes the errors decays cause the massive punt to suffer at long distance layers updates of weights during, think is to suffer at long distance layers. San be reduced by truncating beckmand population. Chasting sight archivation foretain Truncating backmand pergating truncating backmand pergating and the second pergating to the property. Long short-Term memory network: egt I have been in spain for last 10 years. I can speak Awent here, the prediction work depends on sprinformation from previous words. Apoble of learning long-term dependencies by remembering information for lang Perreds is the default behavior.	similarly hi) one to many & iv) many to many.
W.R.T possemeters of network becomes very to grow explorentially as large small or close to zero. This makes the errors decays cause the massive punt to suffer at long distance layers updates of weights during, think is to suffer at long distance layers. San be reduced by truncating beckmand population. Chasting sight archivation foretain Truncating backmand pergating truncating backmand pergating and the second pergating to the property. Long short-Term memory network: egt I have been in spain for last 10 years. I can speak Awent here, the prediction work depends on sprinformation from previous words. Apoble of learning long-term dependencies by remembering information for lang Perreds is the default behavior.	Comage caption) (image caption) (image caption) (image caption) (image caption)
W.R.T possemeters of network becomes very to grow explorentially as large small or close to zero. This makes the earns decays cause the massive punt to suffer at long distance layers updates of weights during, them frocess. San be reduced by truncating becomes layers————————————————————————————————————	Vanishing goodient problem Exploding gradient problem
W.R.T possemeters of network becomes very to grow explorentially as large small or close to zero. This makes the errors decays cause the massive punt to suffer at long distance layers updates of weights during, think is to suffer at long distance layers. San be reduced by truncating beckmand population. Chasting sight archivation foretain Truncating backmand pergating truncating backmand pergating and the second pergating to the property. Long short-Term memory network: egt I have been in spain for last 10 years. I can speak Awent here, the prediction work depends on sprinformation from previous words. Apoble of learning long-term dependencies by remembering information for lang Perreds is the default behavior.	This arises when the gradient of loss function This arises when the slop ten
Small or close to zero. This makes the errors decays cause the massive RNN to suffer at long distance layers updates of weights during. Italian Process. Chasting eight activation function Chasting eight activation function Truncating backward program Truncating backward pr	LIBIT parameters of network becomes very to grow exponentially as large
RNN to suffer at long distance layers. Son be reduced by toursating heteropystion. Choosing eight activation function Choosing eight activation function Truncating backward papaths Trun	small or close to zero. This makes the error decays cause the massive
-) Can be reduced by truncating behavior. -) Can be reduced by charged and characters. -) Can speak fluent -)	PNIN to suffer at long distance layers, in the most property of weights quantity.
Choosing eight activation furched Truncating backward papagation STMN > Long short-Team memory network. Truncating backward papagation egit I have been in spain for last 10 years _ I can speak fluent here, the prediction work depends on parinformation from previous words. Truncating backward papagation egit I can speak fluent egit I have been in spain for last 10 years _ I can speak fluent egit I have been in spain for last 10 years here, the prediction work depends on sprinformation from previous words. Truncating backward papagation for last 10 years _ I can speak fluent egit I can speak fluent egi	-) Can be reduced by tourcoting between population. I -) an be reduced by the chargest
LSTMM > Long short-Term memory network: eg: I have been in spain for last 10 years I can speak fluent here, the prediction work depends on apprinternation from previous words. capable of lowning long-term dependencies by semembering information for lang periods is the default behavior.	chasing sight activation function
here, the psediction work depends on sprinformation from psevious words. To capable of learning long-term dependencies by semembering information for large periods is the default behavior.	etypol nobblid sightlem is maitamilia
here, the psediction work depends on sprinformation from psevious words. Togoble of looming long-term dependencies by semembering information for large periods is the default behavior.	-) LSTMM -> Long short-Team memory network.
-) capable of learning long-term dependencies by semembering information too long perrods is the default behavior.	eg + I hove been in spain for last 10 years - I can to age words.
long periods is the default behavior,	here, the prediction work depends on apprinternation and in the for
long periods is the default behavior,	-) capable of learning long-term dependencies by semembering information
-) RNN : used for Text-to-speech Conversions.	long periods is the default behavior,
	-) RNN: used for Text-to-speech Conversions.

Cympantive Adversariol Network (WAN): Bayes Rule! P(BIA) P(A) P (BIA-) P(A-) 4P (BIA+) P(A+) The generales and the discolarination. P(NB) -> probability of event A occurring given event B has occurred has lugar so milasy Belon anhane saval op 1: P(A) -> Probability of event Anighon in alguar also was P(B) -> lacomination of a serior interest of a real family is the real family is the real family is the discontinuous of the real family is the discontinuous of the real family is the real family in the discontinuous of the real family is the real family in the real family in the real family is the real family in th Gradient Descent algorithmisso styme old restable old along is a popular optamization l'algorithm, used oin zom L & DL to Minimize the cost Hourshop's (linear regrission) updates them? iteratly. 1 starts 25th initial 21 set, of paramenters and evel goodient of cost function. -) It adjust the Henseights in the direction of rolovinioseib bio, solovinioseib =# .) This is very easy to implement. & generated somples. variants of gradient descent (SIGD) probonly long training example is used to compute i) Stochastic gradient descent the gradient and update the parameter at each iteration. (may ked to noisy supdates) i) this-botch GD: small botton of examples are used to compute the graduent and update, the parameter, at each iteration, (better than botch & sup) iii Morentun - based GD: A momentum term is added to gradent update to smooth & the updation process, iv) Adagrad: Adaptive learning note for each parameter based on historical groad: ents

Generative Adversorial Network (GAN): ("GAM is, a deep reasoning model consisting of two neural networks! The generator and the discriminator.

Generator: The generator generates new data similar to the training data. it go takes random noise veator as input and generales a new data sample as outpick. There's to prisided to (1) Discriminator: The discriminator distainguishes between the real and generaled data. It takes do sample as input, and produce probability some data. It takes do sample as input, and produce probability some data. It takes do sample as input, and produce probability some data. It takes do sample as input, and produce probability some data. are trained itagether in an adversarial the generalist Bradischiminator between the generators tries to generate realistic samples bolo tries to correctly colistinguish blus the discriminator, and discriminator the real & generated samples. Applications of GAN: image and video synthesis, data augmentation and il Stockoshie grodient decent slowy brio traiboop out anomaly detection. 1/15/10000 9NY Stugno (Real data samples possemeter. at each iteration. "berten Morenton hased OD: A remember team is added to ment nothing 120000 workely boireday so based setemonal year and steer entires entirely entirely entirely entirely and the books of the stro.bes



Confusion matrix : It is a matrix used to determine the performance the classification model for a given set of test data. Actual values Positive Negative Negative. hollot correct predictions made by the classifue ratio with all other) TP+ FP+FN+TN " Incorrect FP+FN ratio of x correct output provided by correct of promise TP+FP+FN+TN Postio of actual consect output by overall correct predia 2x Recall x Precision + jeul Fiscore is max whohen Pr Recall + Precision FO- 1/1.0) 2.0 L S.1 = . W now with the shold: 1 & leasning note no 03. similar to above process use will go case - by case, in every ase we will get the relimited output: actual output, so these is no weight updation So, w, = 0.7 8 w, -0.6 ose final oseign 13. (1-0) (3) (3) (4)