Maths end-sem (summary)

Unit-1

Summary

- key applications - spam filtering, medical diagnosis, recomm sys.
- core ML task - Association learning, classification, regression & clustering (unsupervised)
- supervised learning - uses labelled data; you define input representation, hypothesis class & version space (subset of hypotheses w/ training data)
- Reinforced learning - (reward feedback)

Unit-2

- Bayesian modelling (probabilistic reasoning) & Gaussian process (function prediction)
- Randomised method introduces stochasticity to improve optimisation
- real world application - image recognition, text generation

Unit-3

- RNN & LSTM used for sequential data like speech
- BPTT trains them by unfolding them over time
- Attention allows to focus on more important words for o/p.

- Memory network & Turing machine enhance reasoning by combining memory & neural nets
- Machine Translation major app. of NLP using DL
- Syntatic & Semantic Parsing helps machine understand grammar rules & sentence meanings.

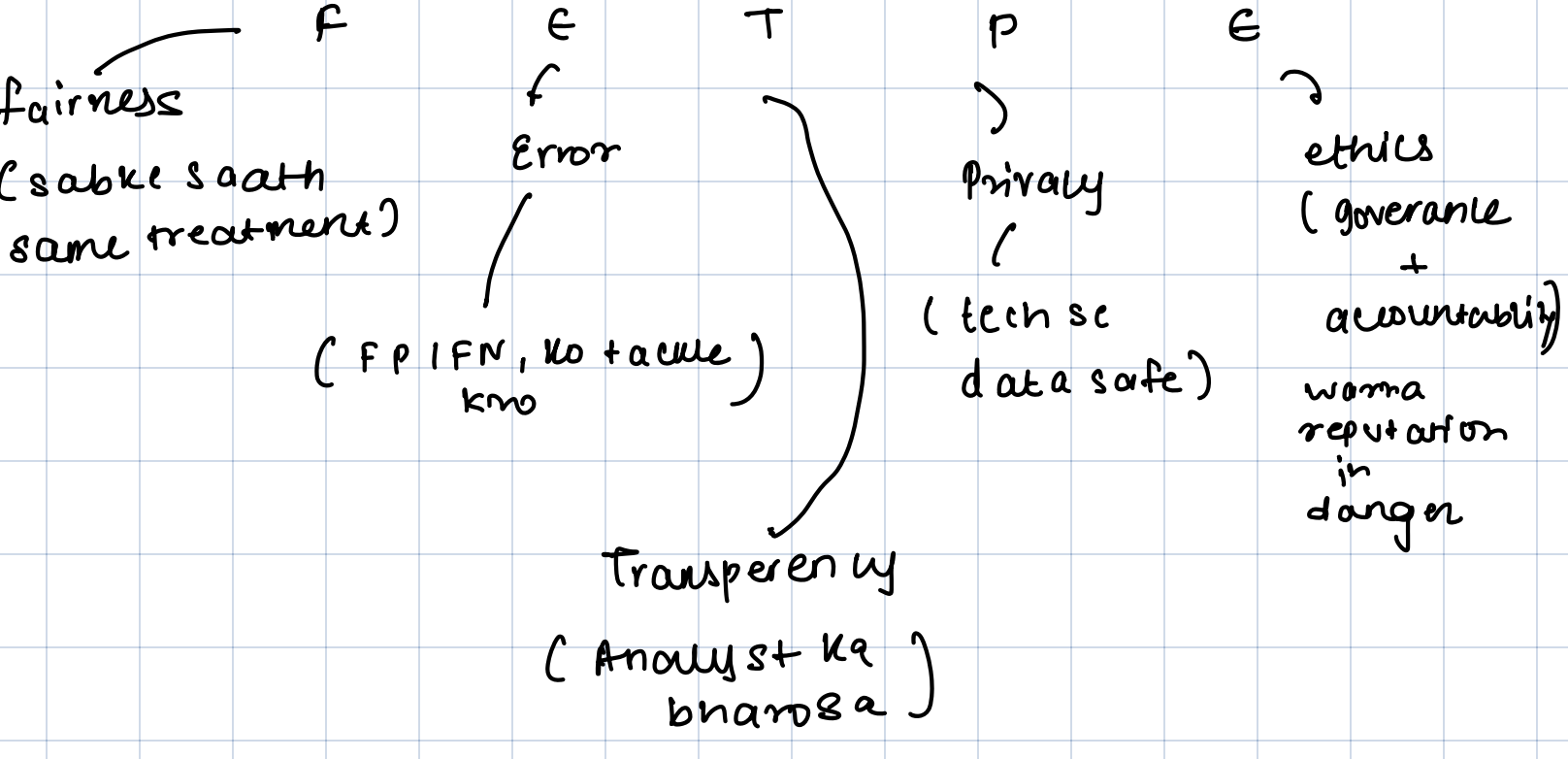
Unit-4

Summary

- Supervised = labelled (spam, phishing)
- unsupervised = unlabelled (Anamoly, UEBA, NIDS & mal. c/w fw)
- Har pipeline will have
  Fetch → Feature → Train - Evaluate → Deploy → Train
- Precision vs recall (Security mein recall important).

Unit-5

(Just insert in any answer).

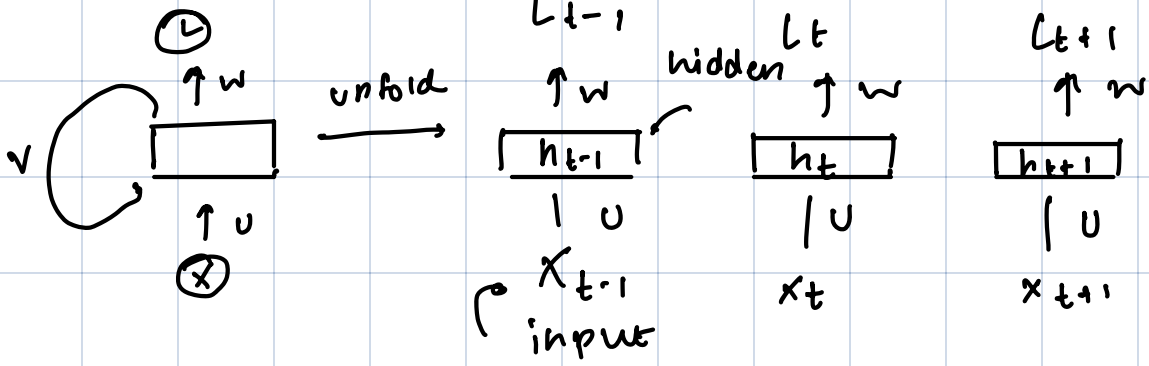


Unit-3

RNN (recurrent neural network)

- It allow network to remember past stuff by feeding the output to the input again

Arch



formula

$$h_t = \tanh(Wx_t + Uh_{t-1} + b)$$

Working..

1) input seq ek-ek karke model mein jate hai.
2) Har step pe model previous step (h_{t-1}) yaad rakta hai.
3) output har time step ka generate hota hai.
4) same weight har step pe use hota hai.

Use cases

- Text prediction
- sentimental analysis
- time series forecasting

pro

- sequential data handle krta hai.
- context maintain karta hai

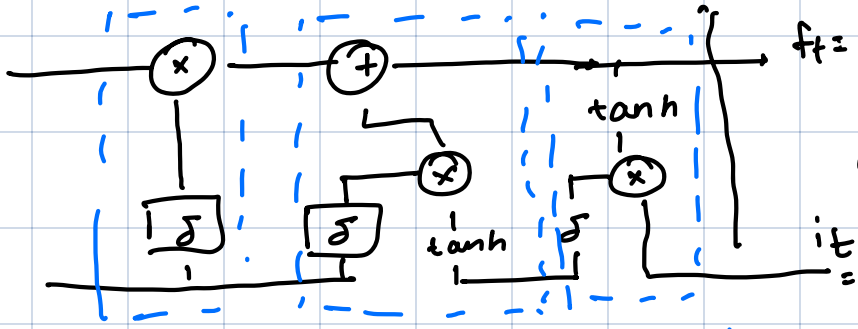
cons

- long term memory nahi hai
- slow training kyuki loop hai.

LSTM (Type of RNN which has long memory).

- Ye vanishing gradient problem solve karta hai using gates

Arch



forget gate
kya bhoolna hai
$f_t = \sigma[W_f x_t + U_f h_{t-1} + b_f]$

input gate
kya naya rakhna hai
$i_t = \sigma[W_i x_t + U_i h_{t-1} + b_i]$

cell state
- jisme long term info store hota hai.

working..

1) input x_t ata hai aur hidden state h_{t-1} ke saath combine hota hai

2) Gate decide krta hai yaad rakhna hai ya bhoolna hai.

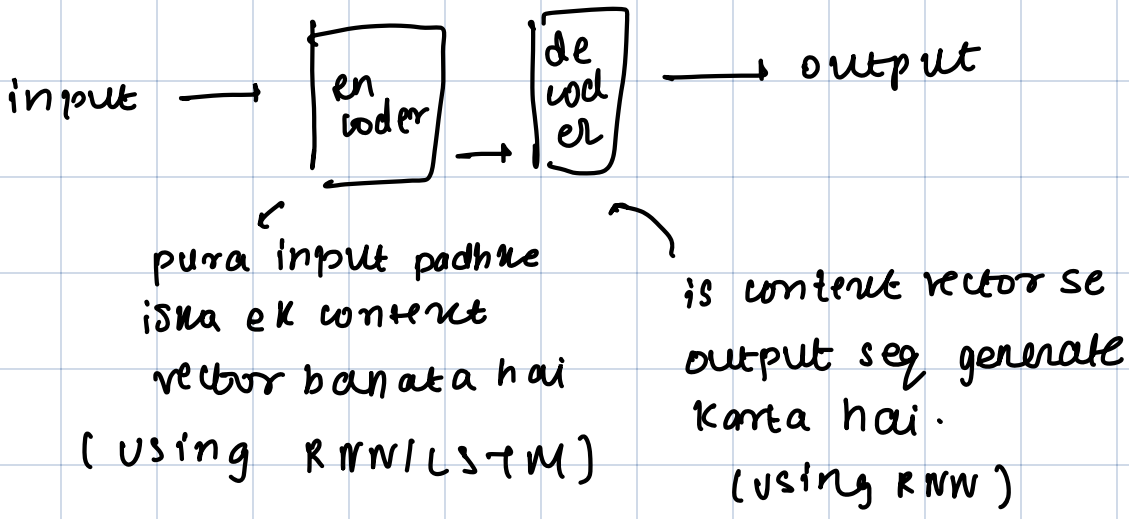
3) Final output ht nikalta hai w/ updated memory.

output gate
kya output dena hai
$O_t = \sigma[W_o x_t + U_o h_{t-1} + b_O]$

Usecase	pro	cons
• long para sentiment analysis • Language translation	• long memory retention • solves vanishing gradient	• slow training • complex

Seq2Seq. (converts input to output like lang. translation)

Arch..



Use case
- chatbot
- machine translation (eng → french)

pro
- Good for variable input & output

cons
- long input = less accuracy ↓ Isko bottleneck bolte hai.

Attention mechanism.. (Solve bottleneck by giving dynamic focus)

Arch
context vector ↑ softmax ↑ Add

working
- o/p token generate krte time i/p words ko dekta hai aur usse importance deta hai

formula
$$AH.(Q,K,V) = \text{softmax}\left(\frac{Q_K^T}{\sqrt{d_k}}\right)V$$

Use case
- Image caption
- speech recog.

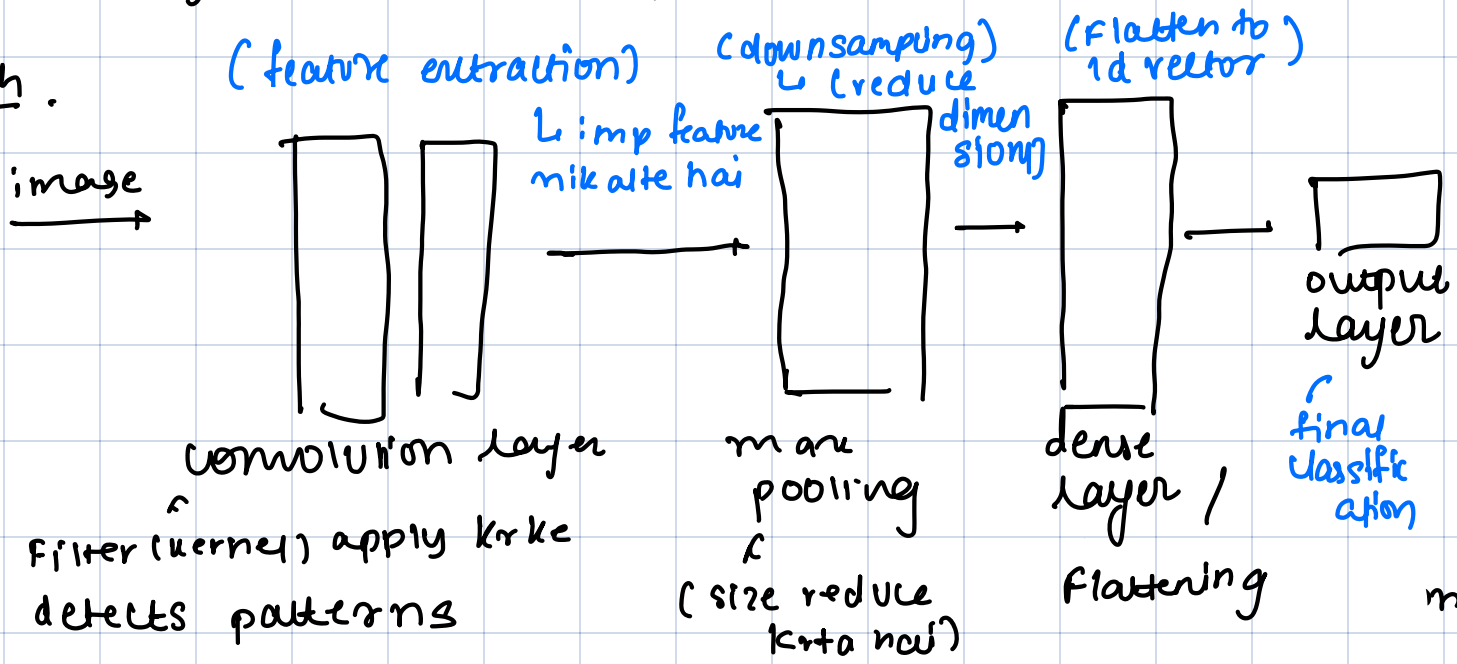
pro
- Better perf on long seq.
- more interpretable o/p

cons
- costly.

Convolutional neural network. (CNN).

Q (Ye images se features (edges, textures, shapes) extract krta hai using convolution filters.

Arch.



Working.

1) i/p image ko filter se scan krta hai
2) ReLU se activation pass hota hai.
3) Pooling layer size choti krti hai
4) final o/p class predict krta hai.

Use case

- Object detection
- Image classification

Pros

- fewer parameters
- efficient image processing

Cons

- not useful for language prob.
- needs large datasets

Unit-4

Internet Arch

App L → Transp L → N/W Layer → link L → Phy L

- App L: end user, eg. HTTP
- Transp L: end2end comm. service. ① TCP ② UDP
- N/W Layer: Internet Layer responsible for routing packet
- link L: handle comm. b/w adj. n/w nodes over phy L
- Phy L: deals w/ physical transmission

key concepts

- packet switching (individually routed for efficient sharing)
- TCP/IP Protocol suite - (standardised way for 2 devices to comm.)
- Distributed Arch - (not centrally controlled)
- client-server model - (client ⇄ (req, resp) server)
- end 2 end principle

Measuring Internet behaviour

- Traffic volume (no. of data thru n/w) byte/sec
- Throughput (successful data transfer)
- Latency (packet source-dest mein delay)
- Packet loss (kitna packet loss hora hai)
- Jitter (packet delay due to fluctuation)
- Flow characteristics (source, dest ke bich mein anamolies)
- Protocol distrib. (Alag Alag protocol ka traffic share) (HTTP, TCP, UDP)
- source/dest analysis (konse ports se source 2 dest.)

Tools for monitoring

- SNMP
- flow monitoring
- Packet sniffing
- Active probing
- Traffic Analysis Tools

Basic idea
(Millions of external & internal data points, humans sab track & co-relate nahi kar sakte . ML real-time main spot karta hai) by spotting patterns

Q1 machine learning application in Network Security.

- Finding threats on a network
  - packets ko monitor karke insider threat ya malware ko find karta hai.
- Safe-Browsing
  - "Malicious" website predict karke block karta hai
- Encrypted traffic malware detect.
  - no need to decrypt every packet
  - packet metadata se anomaly ko detect karta hai.
- End-point malware detection
  - Agar naya malware ho (unseen), usko behaviour se flag karta hai.
- cloud data security
  - suspicious login location + IP check

Basic idea
(Gmail ka spam folder - har din re-train hota to capture new scams like - crypto, air drop).

Q2) Supervised Learning : Spam filtering

Workflow :

Data collection (Lakhon labelled email such as spam / not spam)

↓

Feature Extraction (All caps, domain susp., keywords, no. of Links in mail)

↓

Algorithm used
1) Naive byes - Bayes Theorem use
2) SVM - optimal hyperplane use
3) decision Tree - Tree like structure to answer questions
4) Random F - ensemble method

↓

combines decision trees)

Training / Test (80/20 split, Accuracy Precision- recall test karo)

↓

Classification (Naye unseen emails ko classify karo).

* Benefits - High accuracy, auto updating, no need of manual filter

Basic idea
Bank ka clone page - login cred le lete hain attackers.
(SL learns patterns & features to block new attacks).

Q3) Supervised Learning : Phishing detection.

workflow =

Data collection. [data source. email header, SSL cert, URL det., susp domain / IP. label - phishing / legit]

↓

feature engineering [Gmail - Header - Body - Attachment | website - URL - domain]

↓

model / Algorithm [Gradient Boosting - ensemble that fixes error of previous tree
NN - Model learns complex data to analyse seq.]

↓

model Training / Evaluation [80/20 split. Eval - Tested on test data (never seen before).]

↓

Deployment & monitoring [eg. gateway / extension
continuous monitoring
retrained w/ new data periodically.]

Benefits - High accuracy, proactive detection, scalability, adaptability, feature analysis

challenges. Data quality & Labelling, feature engg complexity, Evolving (new attacks), Class Imbalance, False +ve, Interpretability.

Basic idea
[Agar label nahi mile, UL behaviour sikhta hai fir usse deviations alert karta hai. Better in detecting never seen b4 attacks)

Q.4) Unsupervised Learning : Anomaly Detection

Challenges of Traditional Security

- → Zero day attacks - Never seen b4 attacks
- → sophisticated & evolving threats - Attacker constantly adapt
- → Insider threats - Malicious attacks from within n1w
- → Volume of n1w traffic - Manually defining rules for every attack.

Benefits of Unsupervised learning.

- → Learn Behaviour _ model learns underlying patterns, structure, properties of network traffic.
- → learn deviations - once behaviour is identified, any deviation is flagged as anomaly

Key Algorithms

Clustering (k-means) - (similar flows ko ek cluster Jo bahar nikla wo awara packet) - single Host hitting 1000 ips.

Density (db-scan) - (dense vs sparse region compare) - few packets to odd port 31337)

Dimensionality (PCA / Encoder) - (data / features → compress → suspicious ← error high? ∴ normal data → low dimension) - weird flag combo SYN-FIN

Statistical (one class SVM) - Normal boundary draw - Abnormal login hours outsiders = attack (one)

Time-based. (LSTM) - Expected traffic pattern break - Night time DDOS spike.

workflow..

Data collect (data like logs, packets) → Features (IP, protocol, size, login time, commands, CPU, memory, I/O) → Pre-process (missing, normalise, removing irrelevant feature) → Choose Algo (Algo based on features & expected anomaly) → Train (model learns under patterns & structures normal behaviour) → Score (Apply to unseen data). → Feed back (from security analyst) & reduce false +ve

Benefits - can detect zero-day (naya signature no pblm bhi), Adapt to new threat (behaviour sikta hai), Identification of insider threats (n/w se figure krta hai), Reduce relying on label data (manual labour), discovery of unknown vulnerability

Challenges - High false +ve rate (can flag normal as malicious), Data quality & pre-processing (input data pe depend krta hai noisy, irregular data), Interpretability (anomaly kyu anomaly hai), computational cost (model apne aap sikta hai).

(Anamoly detect + UEBA + network IDS + malware clustering)

Q.5) Unsupervised learning: Practical use cases.

Anamoly detection

→ network ke flow se zero-day attacks pakdo.

data collection → feature → model training → score → Thresholding
(Agar score ↑ flag karo)

Benefits: detection of zero-day, identification of Insider threats, reduced reliance on signature, Adaptable to dynamic env.

UEBA (User & Entity Behaviour Analysis).

→ data logs → Behavioural Profiling → Anamoly score → Risk score → Alert

data logs
(detailed logs of each user
- login/out
- App usage
- N/w, cloud)

Behavioural Profiling
(profile of behaviour of each user)
- cluster
- Association mining
- time series Analysis

Anamoly score
(compared against behaviours)
↓
factors like time, peers, sensitivity

Risk score
(Each user has a risk score)

Alert
(Risk ↑ ↓ Alert)

Benefits: Enhanced insider threat (specifically designed for internal n/w), Identification of comprimised account (Attacker gain access to legit acc), contextual awareness (reduces false +ve), Proactive risk mgmt (identifies risk early)

Network Intrusion detection

→

N/W Traffic → data acquisition (using wireshark)
- Flow based
- Packet based

→ Learning n/w Behaviour
- clustering
- Dimensionality (Deviation → High reconstr. error)
- statistical (Deviation → low probability)

→ detecting anomalous traffic Pattern
- sudden spike
- distant IP / blacklisted
- unusual port / flag
- DoS attack

Benefit -
- detection of zero-day attack (as it focuses on deviation from normal behaviour)
- Identification of stealth attacks (Identifies low volume attacks)
- reduced false +ve (UL)
- Adaptability to n/w change (NIDS adapts over time)

Malware Analysis

→

Sample collection → Grouping similar malware → Malware family & variant identif. → Automated malware classif.
- static (size, timestamp)
- Dynamic (CPU/memory usage)

Grouping similar malware:
- K-means (sim. feature vectors)
- Hierarchical (Hierarchy based on pairwise sim.)
- DBScan (Based on density)

Malware family & variant identif. (malware of same user = same family)

Benefits - scalable, identify new, understanding malware, Improved threat intelligence

Unit-5

Topic	Definition	Purpose	Technique	Tools	Pros
1) Fairness	To avoid algo bias in ML model to ensure fair treatment of all users	To ensure no one is penalised or favored by cybersec	Bias detection, resampling, Fair constraints in loss fn	AIF360, SHAP, LIME	Equal outcome, legal compliance, ethical AI, trust
2) Transperency	Understanding how ML model arive at their decisions understood by a human	To build trust, debugging & ensure regulatory compliance.	Decision tree, SHAP, LIME, saliency map, Rule extraction	SHAP, LIME, TensorBoard	trust, debugging ensures accountability
3) Explainability	Providing understanable reasons behind specific decision made by a ML model.	To explain individual prediction clearly, support investigation & validate model logic	LIME, SHAP, rule extraction, local explainations.	AIx360 interpret ML	Improves decision clarity, validation analysis confidence.
4) Privacy	Protecting sensitive data during model training deployment & interference	To ensure user data is secure, private, & not misuse / leaked.	differential Privacy, Federated learning, Encryption, tokenisation	Crypten, openDP PySyft	Data sec, legal compliance (GDPR, DPDP) user trust
5) Error in ML	mistake made by ML model like false +ve, false -ve, bias etc	To detect, understand & mitigate harmful effects of incorrect predictions.	confusion matrix, retraining, human in loop, monitoring	Scikit-learn, mlflow TensorBoard	maintain perf, avoid breaches ensures trust
6) Ethics in ML	moral principles guiding the responsible use of AI in cybersec	To prevent misuse, ensure fairness, Accountability & protect rights.	ethical audit, Human oversight, red teaming	No dedicated tools.	Trustworthy system, legal protection

Cons	Key Terms	Examples
reduces accuracy. airness is hard to lefine	demographic parity, equal opp. predictive parity	Spam filtering non english mails as spam.
nay expose sensitive nterrals, hard for eep models.	model interpretability post-hoc explain-ation	Explaining why a phishing mail was flagged.
ompu tionally ntensive, not ccurate rep.	local vs global counterfactual shapley value	classifies file as malware due to API calls.
educes utility. curacy · High omplexity & ompute cost.	PII, DPDP act, GDPR	federated learning for GBoard keyboard predictions.
ostly mistake, ifficult to detect arly · may erode rust.	FP, FN, F1-Score, Precision, recall Drift	Model blocks valid user as attacker (False +ve)
ard to enforce volving area	AI act, Bias, accountability, responsible AI.	Unethical monitoring of private msg w/o consent.

Unit-1

ML - Data col → data prep → model selection → training → Evaluation → Deployment

Reinforced learning.

(1) Agent - decision-maker that interacts w/ environm
↳ majdoor

(2) env -
↳
cons stk

(3) state - current situation
↳ 1st floor

(4) Action - decision made by agent
↳ path aar uthaun a

(5) Rewards - feedback by agent after action
↳ dihadi

Markov decision (used to solve reinforced)
↳ Goal - to maximise cummulative reward.

components

① State space - Agent consi situation ya condition mein wo skta hai.

② Action space - action by agent

③ Transition fn - fn that defines probability of going from s to s'.

④ Reward - After action immed. reward.

⑤ discount factor - factor b/w 0 & 1 it represents future reward.

Q - learning (RL algo) - model free

Agent doesn't know external environment.

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max Q(s',a') - Q(s,a) \right]$$

Q(s,a): s = state, a = action; α = learning; r = reward; γ = discount f; max Q(s',a') - Q(s,a)

Challenges of RL

① Exploration vs exploitation - explore kro but exploit bhi kro

② Sample efficiency - acha nahi hai

② delayed reward.

② sparse reward - kaam aaro long time

③ Generalisation - they may struggle to generalise

supervised

① Linear reg
② logistic reg
③ SVM

unsupervised

① K-mean clustering

Appl : Healthcare, Finance & Banking. Retail & commerce, NLP,

challenges: ① Data quality & Availability
② Data privacy & security.

model related to challenges : ① overfitting - model complex (false pattern (noise) features ↑
② underfitting - features ↓ perf ↓ miss kr deta hai (key pattern)
③ model Interpretability
④ scalability.

Algo based challenges : ① Bias / Fairness
② Adversial attack

real-world deploy : ① model deploy / maintence
② Ethical consideration

Hypothesis class

denoted by H
all possible model or f^n that predict o/p.
Algo used to map i/p → o/p.
eg. linear reg, decision tree, Neural nw

version space

subset of hypothesis
↳ shows all correct prediction.

If v.s model struggle

narrow it down to accurately predict the o/p

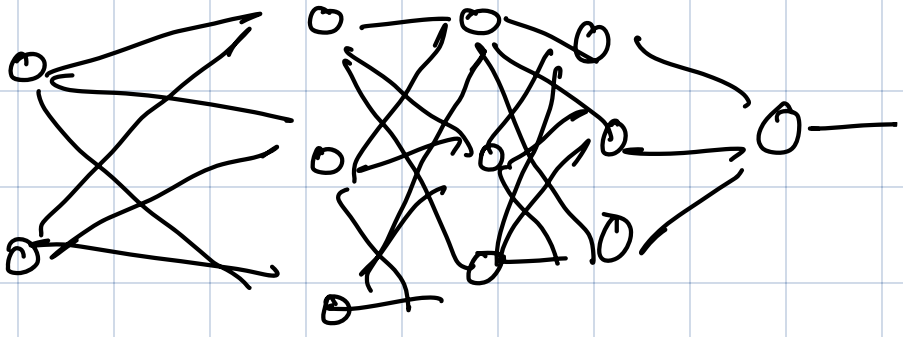
$VS(D) = \{h \in H \mid h(x_i) = y_i \text{ for all } (x_i, y_i) \in D\}$

Unit-2 ..

Deep learning

ANN

- consist of neurons connected by weighted edges
- ilp se learn krke prediction krta hai.



① input

- raw data
- ek neuron ek feature rep krta hai

n feature = n neurons.

② Hidden

- Intermediate calc
- recieve i/p. apply weighted sum & passes thru activation fn.

③ O/p (final)

no. of neuron depend on task

sigmoid :- $y = \frac{1}{1+e^{-z}}$

- no. of HL & neurons in it depends on complexity

$z = wx + b$ — bias
input
weight sum before activation
weight

$a = f(z)$

Activation function.

① sigmoid - Binary classif.

② ReLU - Rectified Linear Unit

$f(z) = \max(0, z)$

③ tanh :- $f(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$

center the o/p around 0.
making optimisation easier.

④ softmax

converts logits into probab. for multiclass problem.

Regularisation

- Improve generalisation.
- Accuracy ↑
- complexity ↓
- prevent overfitting.
- reduce weight of neuron.

CNN (used in img classification, facial recog, comp vision)

CNN	RNN
# spatial data	* sequential
# parallel processing	* sequential processing
. feed forward	. feedback loop / hidden arch
. no mem of previous	. memory of previous via hidden
. fast	. slow
. min vanishing gradient	. common vanishing
. pre-trained CNN available	. fewer model available

Variational Autoencoder

(combine neural nw w/ probabilistic modeling.

① Encoder :-

- map i/p data x to latent space
- encoder outputs two vectors instead of one vector
  * mean (μ) * standard deviation (σ)

② Sampling (latent space) :-

- $z = \mu + \sigma \cdot \epsilon \quad \epsilon \sim N(0,1)$

Allows gradients to pass through back propogation.

③ Decoder

maps vector z back to the original data space producing reconstruction.

VAE loss fn :. diff b/w i/p & reconst o/p.

$$L_{VAE} = E[\|x - \hat{x}\|^2] + D_{KL}[q(z|x)\|P(z)]$$

Adv

- can generate new sample from latent space
- useful unsupervised & data comp.
- clear probabilistic interpretation

Applications

- data compression
- anomaly det.
- Image generation