Agenda.

1) Ethical Guidelines

① **Topic:** Latent Vector Vs Embedding Vector

② Synthetic Data vs Augmented Data

③ GAN Vs Diffusion Model

Generative Vs Descriminative Model

Over parametrized Vs Regularization

→ Zero Shot Vs Few shot learning

Fine tuning vs Transfer learning

→ Attention Mechanism Vs Self Attention

Tokenization vs Embedding

Perplexity Vs Accuracy

→ Recall Vs precision

Overfitting Vs Underfitting

Epoch vs Iteration vs Batch

Noise vs Signal

→ Promp tuning Vs promting Vs

Prompt Injection

→ Intelligence Explosion

→ Federated Learning vs Explainable XAI

→ Black box vs White Box

→ Edge AI vs Style Transfer

20 → Regularization vs Generalization

Ethical Guidelines → AI / Robots / →

↓

AI Model → Set of Rules

↓

Responsible AI

←

Ethical

1) Transparency.

2) Data privacy.

3) Bias / Fairness

4) Misinformation Control → Hallucination

5) Accountability.

6) Environment Friendly. → Resource

7) Consent & Right

8) Accessibility

10) Continuous Monitoring

⊛ Latent Vector Vs Embedded Vector

Latent Vector
↓
Model Training
↓
Set → parameter
↓
Latent → Vector
↓
Lower Dimensional

Embedded Vector
↓
Word / Text → Data
↓
Training
↓
Vector – embed vector
↓
Pretrained vector

(2) → Synthetic Data
↓

Augmented Data
↓

Data Generated By
Model → GAN
is known Synthetic Data

Data Augmentation
↓
1) Scaling
2) Rotating
3) Color / Angle
↓
Basic Remove
↓
Augmented Data

③ → Pretraining Vs Finetuning

Data → start (Training)
↓
Data
↓
Pre training.

Model + Dataset
↓
Training
↓
Fine tuning

(4) → Over parametrized
↓
→ More Number params
→ Extra Training

Regularization
↓
Techique to
Prevent Overfitting
↓

↓ U

Overfit

1) Dropout
2) early stopping

⑤. GAN

Generative Adversial Network

↓

Generatiεo   Disrimator
                     ↓
Image → copy   Real/ Fake

↓

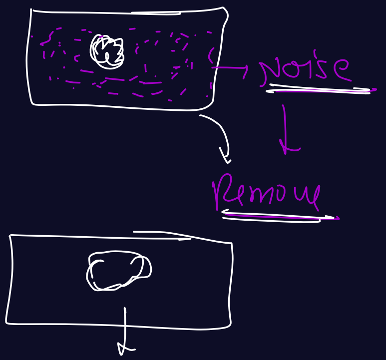
(Images) → Syntentic Data

Diffusion Model

↓

Images   (Dalle)

↓



↓

Image.

⑦ → Gradient Descent

↓

Model optimsation

↓

Redue the Error to

Global Minima

① * Expensive

Sochastic Gradient Descent.

↓

Singl data

↓

optimation Algorthm

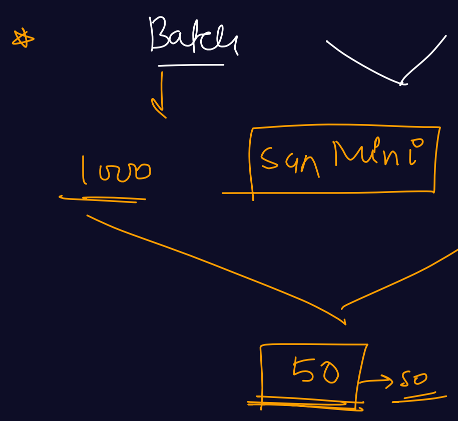
↓

⊛ Time

② → Slow

② Accuray

SGD Min Batch



⑫

⑧ . Embedding

↓

② Token → Numerical

↓

Embedding vector

①

Tokenzation .

↓

① Sentence → words

↓

Token

⑨ . Perplexity

↓

How well Model is

⊛ Accuray .

↓

Prediction = 5

Predicting the Next word

How are correct from 5

↓ ① — Pm
5

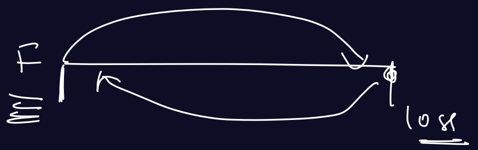
⑩ . Overfitting
↓
रटट
⇓

{ Training Data — ✓
{ Testing Data → ✗ (pstana)
↓
Overfitting .

Underfitting
↓
नालायक
⇓

Traing Data → Test 9n ✗
Testing Data — ✗
↓
Underfitting

⑪ → 1) Epoch → One complete iteration



{ — epoch

(2) → Batch → subset of dataset → that we are passing to Model -

1000 row

{50} ⇒ Batch .

(3) → Iteration ⇒ No of Batch to pass all data =

$$\frac{\overset{20}{\cancel{1000}}}{\cancel{50}} = \underline{\underline{200}}$$

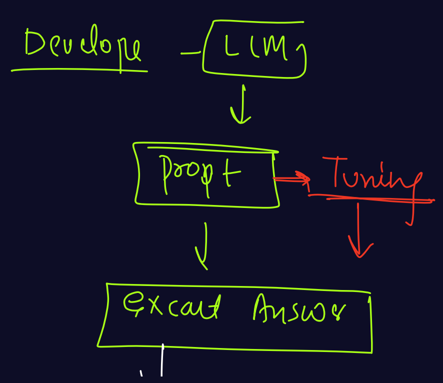
→ Noise Vs Signal ⇒

↓

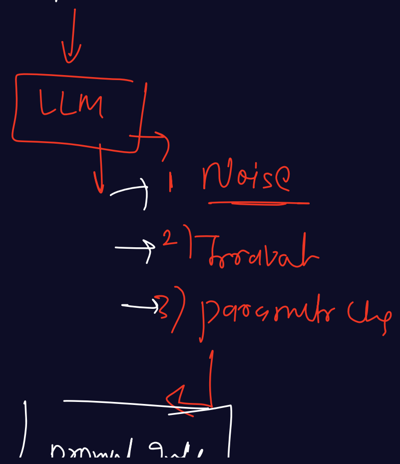
① Irrelavant pattrn / Irrelvant data → Noise

② Signal : data that convey Meaningful Information

⊛ promt tuning *



Prompt Injection



↓

Prompt tuning

↓

Decrease ↓

⊛ Black Box AI

↓

No Transparency

↓

Not able to get Source of Answer

↓

Transparent ✗

White Box AI

↓

Transparent Model

↓

You can get the Resource

⊛ Search GPT → web

⊛ Device ⇒ Data

Edge AI

↓

Meta → Llama

GPT —

↓

⊛ Mobile AI Model Edge

Mobile → edge M

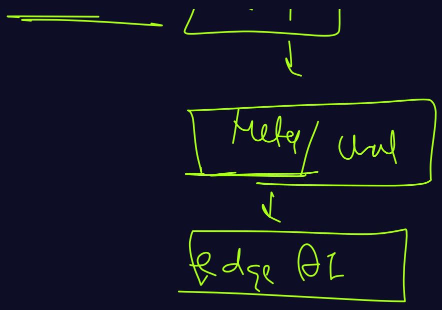
⊛ Explainable AI

↓

You can understand

↓

Decision → Making


Helper/ Util
Edge AI