① » Machine learning is designing a model, which can learn and make predictions from previous data.

Eg:- Self driving cars., Emails, recommendations systems, etc.

② Every business nowadays needs machine learning bcz it helps convert large amt of data into insights quickly & efficently.

③ Types
- Supervised ML.
- Unsupervised ML.
- Reinforcement ML.

3a Supervised ML → Also called predictive learning approach.

i) Steps:
- Labels are given.
- then labeled (sorted) data is given.
- when Model is done training, test data is given, to make predictions.

ii) Methods
- Classification. (categorical)
  - Multiclass.
  - Binary. (only 2. Eg:- 1 & 0).
- Regression. (numerical)
  - Linear Regression
  - Non-linear Regression.
    - Quadratic
    - Cubic
    - Multiple
    - Polynomial.

Regression Equations.
- Linear ($f(x) = y = mx + c$)
- Quad ($y = mx^2 + m'x + c$)
- Cubic ....
- Polynomial ($y = m_1x^n + m_2x^{n-1} \ldots + b$)
- Multiple ($y = b + m_1x_1 + m_2x_2 \ldots m_nx_n$).

(3b) Unsupervised learning.

(i) → Goal is to find interesting patterns in the data. Sometimes called knowledge discovery.

∴ Explores pattern & predict outputs by forming clusters.

(ii) Methods
- → Clustering (Group similar data together).
- → Association (if one data dependent on other) [Bread & Butter].

Clustering can be done by shape, colour or characters.

(3c) Reinforcement Learning.

(i) → Used for learning how to act or behave when occasionally given reward or punishing signals.

Simply put it follows a trial and error method.

(ii) Dataset in unsupervised is not labelled, reinforcement learning doesn't have a predefined dataset.

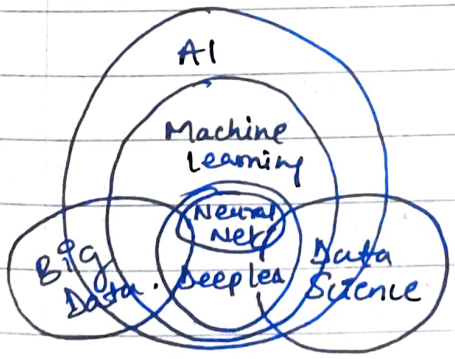
(iii) Methods → Exploitation.
→ Exploration.

(4) Differences b/n learnings

Aim → SML → Calculate outcomes [weather, risk].
→ US ML → Discover pattern [Recommendation in streaming, Anomaly detection].
→ R ML → Learn a series of action [self driving cars, Gaming].

Algorithms → Linear Reg, Logistic Reg, SVM, KNN, etc.
→ K-Means, C-Means, Apriori
→ Q-learning, SARSA.

(5) Machine Learning Terminologies.



(5a)

AI ⇒ where human intelligence is incorporated into machines that self-learn based on algorithms.

ML ⇒ Part of AI, that uses statistical learning algorithms to build machin that have ability to learn & improve.

Deep learning → Uses neural networks to build machines that can filter input data to predict & classify information.

Neural Networks → Heart of deep learning, that mimic the operations of human brain to recognize relationship b/n data.

(5b) Algorithms VS models

⇒ Algos are programs (math + logic) that adjust themselves to perform better as data changes.

⇒ A ML model is a file that has been trained to recognize certain types of pattern.

∴ A train a model over set of data using algorithms. Model uses algo to reason with & learn from data.

(5c) Steps in Machine Learning

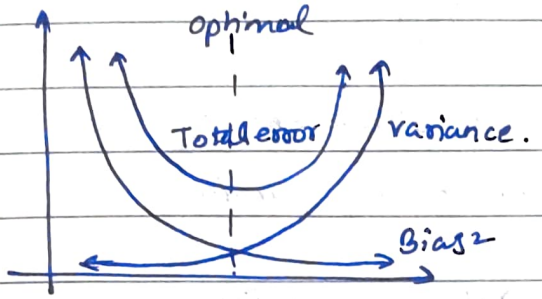
Gather data → Prepare data → Choose a model → Train the model → Evaluation → Hyperparameter tuning [choosing parameters with least errors] → Prediction

(5d) Data split into 2 parts.
- → 25-30% (Testing data)
- → 70-75% (Training data).

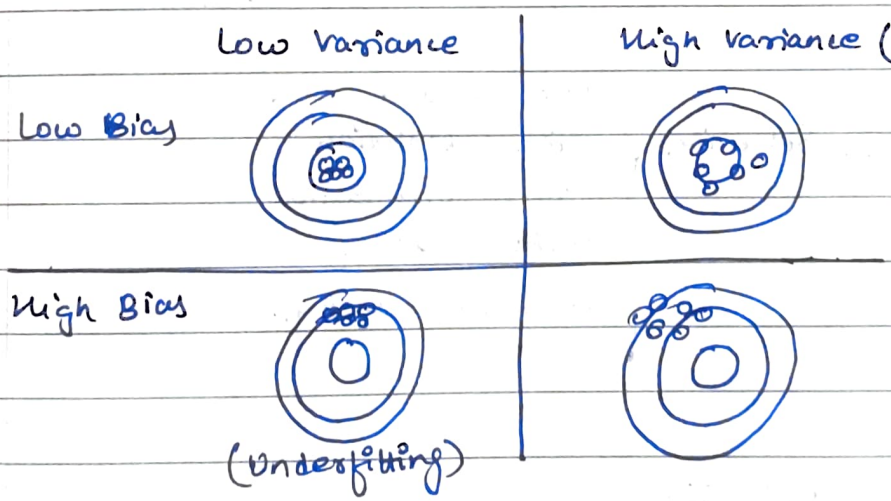
5e) Bias & Variance

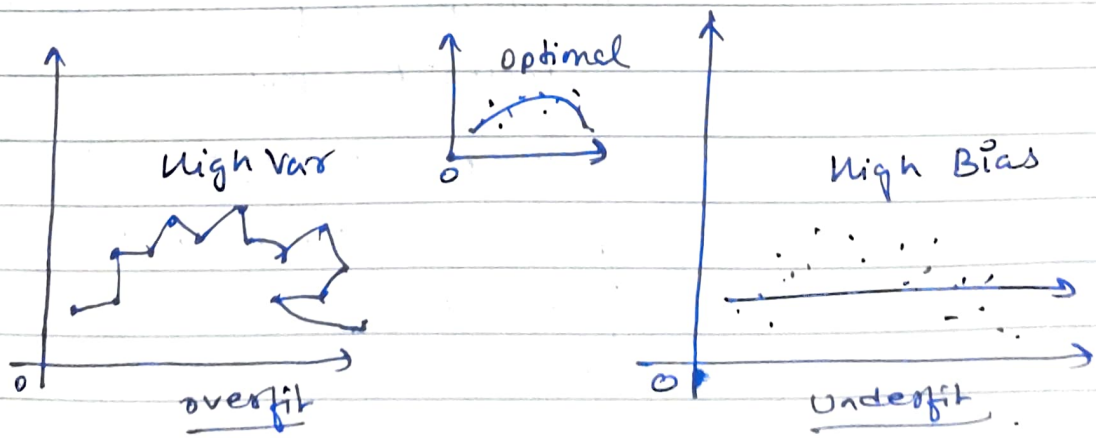
Bias → Difference b/n predicted and actual value.

Variance → Difference b/n fits of dataset i.e training & testing data.



As variance ↓ Bias ↑
As Bias ↓ Variance ↑
∴ A good balance b/n variance & Bias is necessary.





(6) Curse of dimensionality.

(6a) Dimensions → No: of attributes a dataset has.

Eg:- Dataset of a person X.

Height, weight, BMI } 3 attributes, No: of dim ⇒ 3.

(6b) Curse of Dim means as dimensions of the dataset increase, the errors increase. & runtime becomes expo.

(6c) Dimensionality Reduction techniques
- → Feature Selection
- → Feature Extraction

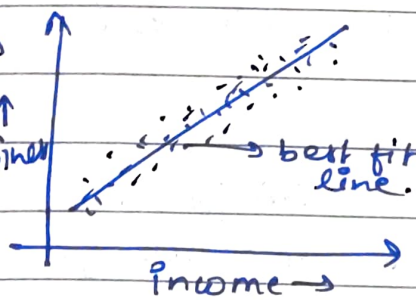
Feature Selection.
- Low variance filter.
- High correlation
- Multi collinearity
- Feature ranking
- forward Selection.

Feature Extraction
- Principal component Analysis
- Factor Analysis
- Independent Component Analysis
- Linear Discriminant Analysis
- Quadratic.

(7) Linear Regression

(7a) ⇒ A method in supervised learning, used for finding linear relationship b/n target and one or more predictors.

(7b) ⇒



⇒ $y = mx + c$

↳ equation.

(8) Testing Cross-Validation

(8a) ⇒ We have already seen how dataset is divided into testing & training data. (Validation dataset can also be taken out).

(8b) In cross-validation test data is exchange for equal training data & model is checked many times.

(8c) Particularly used when data is limited.

(8d) Types
- Holdout method.
- K-fold.
- Stratified K-Fold.
- Leave P-Out.

(9) Regulations (Regularization)

(9a) ⇒ Basic terms ⇒ Optimization → deals with finding optimal solution.

Objective func → Any fn optimized during training is an objective fn.

Cost/Loss/Error fn → When we are minimizing the fn, it is known as cost or loss fn.

→ When maximizing it is called reward fn.

* Loss func determine error b/n predicted & targeted value and cost func is avg loss over the entire training dataset.

(9b) Regularization is a technique used to reduce errors by fitting the func approx on given Training set and avoid overfitting.

Adds info to the coefficients. [Cost fn = Loss fn + Regularization term].

(9c) Types → L1 (Lasso regression)
→ L2 (Ridge regression).

(10) Learning curve → Shows how error goes ↑ or ↓ as training set ↑

Used to compare algorithms, model parameters and determine the amt of data used for optimized result.

Types → Optimized learning curves. (when model being optimized).
→ Performance learning curves. (Curve when model evaluation & selection).

(11) Classification

(11a) An supervised algorithm that deals with identifying the class to which an instance belongs.
[category]

(11b) Types → Binary [0 or 1] [True or False]
→ Multi class [0, 1 ... n]

$y = f(x)$

(11c) Linear models → Logistic regression, Support vector machines.

Non-linear models → KNN, SVM Kernel, Naive Bayes, Decision & Random forest trees.

(11d) Metrics → Logg loss (or) Cross entropy loss.
→ Confusion matrix.
→ AUC-ROC curve.

(12) Error & Noise.

⇒ In ML dataset errors are known as noise.

⇒ Types:
- Attribute Noise. (Missing, Don't care, erroneous values)
- class noise. (Contradictory values, Mislabeled examples.)

(13) ⇒ Parametric vs Non-Parametric models

⇒ Model with fixed no: of parameters is parametric. & vice versa. (parameters ↑ data ↑)

⇒ Eg:- Linear regression, Linear support vector machines, Logistic regression, Naive Bayes, Perceptron. (SVM).

of parametric

Non-param → KNN, Decision trees, SVM with Kernels, ANN.

Parametric.	Non-parametric.
Homogeneous	Homogenous & Heterogeneous
Ratio or interval	Ordinal or nominal.
Mean used	Median used
More conclusions	Less affected by outliers.