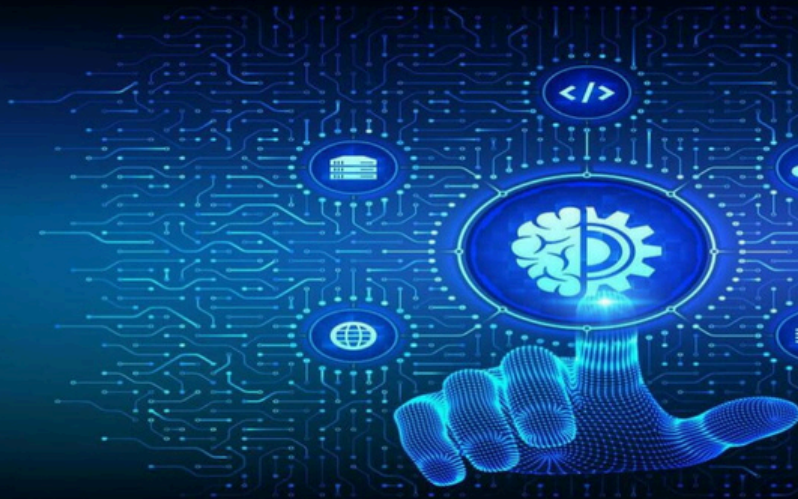


MACHINE LEARNING

VECTOR ILLUSTRATION



অনেকেই মেশিন লার্নিং শেখে, কিন্তু নিচের মূল বিষয়গুলো না জানলে শেখাটাই অন্ধকারে ঢিল ছোঁড়ার মতো হয়। শুধু কোড রান করে বা মডেল বানিয়ে কোনো লাভ হয় না, যদি ভিতরের "বুদ্ধিটা" বোঝা না যায়।

মেশিন লার্নিং শেখার জন্য যে বিষয়গুলো না জানলে শেখা না শেখা সমান (সহজ ভাষায় ব্যাখ্যা)

1. Supervised vs Unsupervised Learning

□ কেন জরুরি:

আপনি কী শিখছেন সেটাই যদি না জানেন, তাহলে সামনে যাবেন কিভাবে?

জানতে হবে:

- Labeled Data মানে কী
- Classification vs Regression
- Clustering কবে লাগে

2. Train/Test Split & Model Evaluation

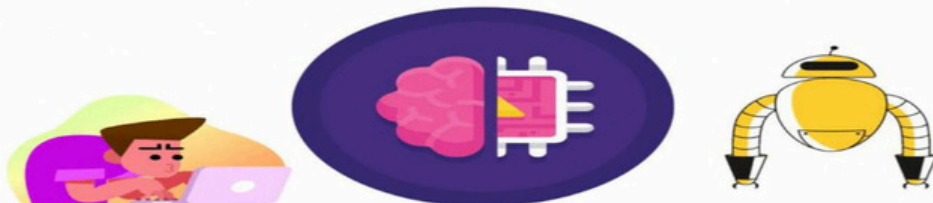
□ কেন জরুরি:

মডেল ঠিকভাবে কাজ করছে কিনা, সেটা না জানলে আপনি ভুল মডেলেই ভরসা করবেন।

জানতে হবে:

- Train/Test/Validation Split
- Accuracy, Precision, Recall, F1-score
- Confusion Matrix

Machine Learning Statistics



3. Bias-Variance Tradeoff

□ কেন জরুরি:

Overfitting vs Underfitting কেন হয়, সেটা না বুঝলে আপনি মডেল "শেখাবেন" কীভাবে?

জানতে হবে:

- High Bias মানে কি
- High Variance মানে কি
- কিভাবে ব্যালান্স করতে হয়

4. Feature Engineering

□ কেন জরুরি:

Raw data কখনোই সরাসরি মডেলে ঢুকানো যায় না — বোঝা দরকার কোন ফিচার দরকার, কোনটা না।

জানতে হবে:

- Feature Selection
- One-Hot Encoding
- Normalization vs Standardization

5. Algorithms এর ভিতরের লজিক

□ কেন জরুরি:

কোন সমস্যায় কোন অ্যালগরিদম ভালো, তা না জানলে “try and error” ছাড়া উপায় থাকে না।

অবশ্যই শিখতে হবে:

- Linear Regression
- Logistic Regression
- Decision Tree
- KNN
- SVM
- Naive Bayes
- Random Forest
- Gradient Boosting (e.g., XGBoost)

6. Cross Validation

□ কেন জরুরি:

একবার Train/Test ভাগ করে ফলাফল ভাল পেলেই কাজ শেষ না — CV না জানলে ফলাফল ভুল হতে পারে।

জানতে হবে:

- K-Fold Cross Validation
- Stratified Sampling

7. Hyperparameter Tuning

□ কেন জরুরি:

মডেলের পারফরম্যান্স বাড়ানোর জন্য খুবই জরুরি — অনেকেই এটাকে বাদ দেয়!

জানতে হবে:

- Grid Search
- Random Search
- Learning Rate, Depth, Estimators

8. Dimensionality Reduction

□ কেন জরুরি:

যখন ফিচার অনেক বেশি, তখন noise কমিয়ে আসল ফিচার বের করা লাগে।

জানতে হবে:

- PCA (Principal Component Analysis)
- t-SNE (visualization purpose)

9. ML Workflow ও Pipeline

□ কেন জরুরি:

একা মডেল বানালেই হয় না, পুরো process জানা দরকার।

জানতে হবে:

- Data → Clean → Feature → Model → Evaluate → Improve → Deploy

সংক্ষেপে মনে রাখার মন্ত্র:

"Data বুঝো + Algorithm জানো + Evaluate করো + Improve করো = Real Machine Learning"

Bonus: না জানলে সময় নষ্ট এমন কিছু ভুল:

- শুধুই Accuracy দেখলে চলবে না
- Data না দেখে মডেল চালানো
- Cross-validation বাদ দেওয়া
- সব সমস্যা সমাধানে Random Forest লাগানো

