Pratham Chandratre

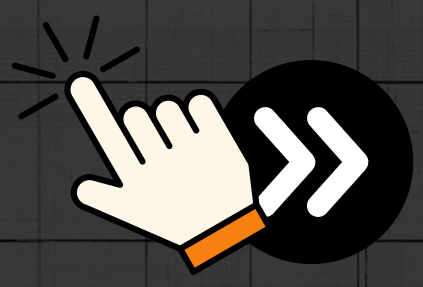
AI/ML Engineer

100

AI Terms

Every Data Professional should Know

Swipe for more



Pratham Chandratre
AI/ML Engineer

AI Fundamentals

- **Algorithm: Step-by-step procedure to solve a problem.**

- **Artificial Intelligence(AI): Machines mimicking human intelligence.**

- **Neural Network: Brain-inspired interconnected nodes for pattern recognition.**

- **Supervised Learning: Training models with labeled data (e.g., spam detection).**

- **Unsupervised Learning: Finding patterns in unlabeled data (e.g., clustering).**

- **Reinforcement Learning: Learning via rewards/penalties (e.g., game AI).**

- **Bias-Variance Tradeoff: Balancing underfitting and overfitting.**

- **Feature Engineering: Crafting input variables to improve models.**

- **Hyperparameter: Configurations set before training (e.g., learning rate).**

- **Loss Function: Measures model prediction error (e.g., MSE).**

Pratham Chandratre
AI/ML Engineer

AI Fundamentals

Overfitting: Model memorizes training data, fails on new data.

Generalization: Model performs well on unseen data.

Epoch: One full pass through the training dataset.

Activation Function: Adds non-linearity to neural networks (e.g., ReLU).

Gradient Descent: Optimizes models by minimizing loss.

Backpropagation: Adjusts neural network weights via error correction.

Ensemble Learning: Combines models for better accuracy (e.g., Random Forest).

Transfer Learning: Reusing pre-trained models for new tasks.

Model Drift: Performance decline due to data changes over time.

Explainable AI (XAI): Making model decisions interpretable.

Pratham Chandratre
AI/ML Engineer

Machine Learning

- Regression: Predicts continuous values (e.g., house prices).
- Classification: Assigns labels (e.g., fraud/not fraud).
- Clustering: Groups similar data points (e.g., customer segments).
- Decision Tree: Flowchart-like model for decisions.
- Random Forest: Ensemble of decision trees for robustness.
- Gradient Boosting: Sequentially corrects errors (e.g., XGBoost).
- Cross-Validation: Tests model generalizability (e.g., k-fold).
- Precision/Recall: Metrics for classification performance.
- ROC-AUC: Evaluates classifier tradeoffs (TPR vs. FPR).
- Confusion Matrix: Tracks TP, TN, FP, FN.

Pratham Chandratre
AI/ML Engineer

Machine Learning

- **SVM (Support Vector Machine): Finds optimal hyperplanes for classification.**
- **K-Means: Groups data into k clusters.**
- **PCA (Principal Component Analysis): Reduces dimensionality while preserving variance.**
- **Regularization: Prevents overfitting (e.g., L1/L2).**
- **Outlier Detection: Identifies anomalies in data.**
- **Feature Selection: Picks most relevant input variables.**
- **Dimensionality Reduction: Simplifies data without losing critical info.**
- **Bagging: Trains models in parallel (e.g., Random Forest).**
- **Boosting: Trains models sequentially to fix errors (e.g., AdaBoost).**
- **MLOps: Automates ML lifecycle (deployment, monitoring).**

Pratham Chandratre
AI/ML Engineer

Deep Learning

- **CNN (Convolutional Neural Network): Processes grid-like data (e.g., images).**
- **RNN (Recurrent Neural Network): Handles sequences (e.g., time-series).**
- **Transformer: Architecture for parallel processing (e.g., BERT, GPT).**
- **Autoencoder: Compresses data, detects anomalies.**
- **GAN (Generative Adversarial Network): Generates synthetic data (e.g., deepfakes).**
- **Attention Mechanism: Focuses on relevant input parts (e.g., translation).**
- **Dropout: Reduces overfitting by deactivating neurons randomly.**
- **Batch Normalization: Stabilizes training by normalizing layer inputs.**
- **Learning Rate: Step size in gradient descent.**
- **LSTM (Long Short-Term Memory): RNN variant for long-term dependencies.**

Pratham Chandratre
AI/ML Engineer

Deep Learning

- GRU (Gated Recurrent Unit): Simplified LSTM for efficiency.
- Embedding Layer: Converts categorical data to vectors.
- Vanishing Gradient: Issue where gradients become too small during training.
- YOLO (You Only Look Once): Real-time object detection model.
- Fine-Tuning: Adapts pre-trained models to new tasks.
- Tokenization: Splits text into words/subwords.
- Beam Search: Heuristic for sequence prediction (e.g., text generation).
- Transformer Architecture: Uses self-attention for NLP tasks.
- Self-Attention: Weights input relevance dynamically.
- Weight Initialization: Method to set initial neural network weights.

Pratham Chandratre
AI/ML Engineer

Natural Language Processing

- **Tokenization: Splitting text into words/subwords.**
- **BERT: Bidirectional language model for context understanding.**
- **TF-IDF: Scores word importance in documents.**
- **Word2Vec: Generates word embeddings for semantic analysis.**
- **Sentiment Analysis: Detects emotion in text (e.g., reviews).**
- **NER (Named Entity Recognition): Identifies entities (e.g., people, places).**
- **Stemming/Lemmatization: Reduces words to root forms.**
- **Seq2Seq: Converts input sequences to output sequences (e.g., translation).**
- **Masked Language Modeling: Predicts masked words in sentences.**
- **GPT: Generative language model for text creation.**

Pratham Chandratre
AI/ML Engineer

Natural Language Processing

- Zero-Shot Learning: Predicts unseen classes without training.
- Text Summarization: Condenses long text into key points.
- Topic Modeling: Discovers themes in documents (e.g., LDA).
- POS Tagging: Labels words with grammatical roles (e.g., noun, verb).
- Dependency Parsing: Analyzes grammatical structure.
- Corpus: Large text dataset for training/models.
- Stop Words: Common words excluded from analysis (e.g., "the").
- N-gram: Contiguous sequence of n words.
- Token Embeddings: Numerical representations of tokens.
- BPE (Byte-Pair Encoding): Subword tokenization method.

Pratham Chandratre
AI/ML Engineer

Ethics & Governance

- **Bias: Skewed model outputs due to unfair data.**
- **Fairness: Ensuring equitable outcomes across groups.**
- **Explainability: Making AI decisions interpretable.**
- **GDPR: EU regulation protecting user data privacy.**
- **Adversarial Attack: Manipulates inputs to fool models.**
- **Data Anonymization: Protects user identities.**
- **Model Audit: Evaluates models for compliance/ethics.**
- **AI Governance: Policies for responsible AI use.**
- **Transparency: Clear disclosure of AI decision-making.**
- **Responsible AI: Ethical design/deployment of AI systems.**

Pratham Chandratre
AI/ML Engineer

Ethics & Governance

- **Privacy-Preserving AI: Techniques to protect data (e.g., federated learning).**
- **Federated Learning: Trains models on decentralized data.**
- **Model Cards: Documents model performance/limitations.**
- **Ethical AI: Aligns AI with moral principles.**
- **Data Sovereignty: Laws governing data storage by geography.**
- **Algorithmic Accountability: Responsibility for AI outcomes.**
- **AI Safety: Ensures AI systems behave as intended.**
- **Synthetic Data: Artificially generated data for training.**
- **Differential Privacy: Protects individual data in datasets.**
- **Compliance: Adherence to legal/regulatory standards.**

Pratham Chandratre

AI/ML Engineer

If you find this helpful, please like and repost it with your friends