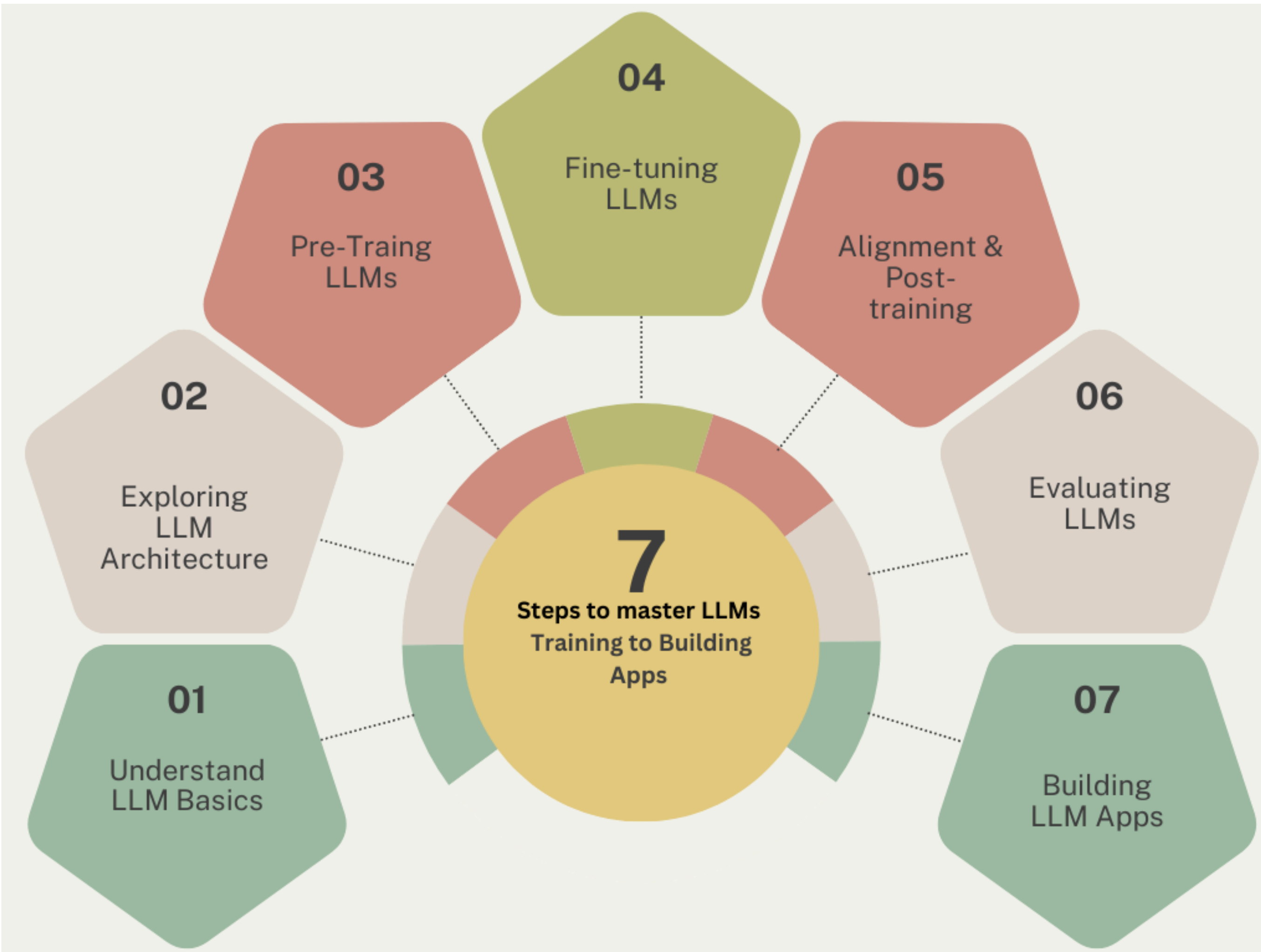
Fine-Tuning Large Language Models - Key Concepts and terms



Category	Details
LLM Architectures	Knowledge of established models (GPT-3, BERT, T5, etc.) and recent developments like Jurassic-1 Jumbo, Megatron-Turing NLG, Pathways Language Model, LLaMA2, Mixtral, GPT-4, and Bard.
Data Preparation	- Data Collection: Domain-specific data - Data Cleaning: Noise removal - Data Augmentation: Synonym replacement, back-translation - Tokenization: Proper text tokenization - Emphasize focus on domain-specific datasets and techniques like data balancing for skewed datasets.
Preprocessing Techniques	- Normalization: Consistent text format - Handling Special Tokens: Start/end tokens, padding, truncation - Sentence Detection and Error Correction
Model Configuration	- Hyperparameter Tuning: Learning rate, batch size, sequence length, epochs - Optimizer Selection: Adam, AdamW, PPO - Scheduler: Learning rate schedulers - Techniques like Gradient Clipping and Mixed Precision Training for efficiency.

Training Process	- Loss Functions: Cross-entropy, etc. - Regularization: Dropout, weight decay - Gradient Accumulation: Large batch handling - Sparse Attention: Reduces computational cost by focusing on relevant parts of the input. - Mixture of Experts (MoE): Trains a group of smaller experts for different subtasks. - Proximal Policy Optimization (PPO): A reinforcement learning algorithm for training models with complex policies, potentially useful for interactive LLM tasks.
Transfer Learning	- Feature Extraction - Fine-Tuning vs. Feature-Based Transfer Learning - Prompt Engineering: Crafting effective prompts to guide the model for specific tasks. - RLHF (Reinforcement Learning from Human Feedback): Improves model performance through human interaction.
Advanced Fine-Tuning Techniques	- Task-Specific Heads - Unfreezing Layers - Layer-Wise Learning Rate Decay (LLRD) - Differential Learning Rates - Multitask Fine-Tuning-Task-Agnostic Meta-Learning: Adapts the model to new tasks quickly with minimal fine-tuning. Pretraining Objectives:Include details on objectives like masked language modeling (MLM), causal language modeling (CLM) - Continuous Adaptation with Pre-training (CAPT): Enables continual learning on new data streams. - Direct Preference Optimization (DPO): A recent advancement in policy optimization algorithms that might be applicable for fine-tuning

	LLMs in dynamic environments. - Low-Rank Adaptation (LoRA),QLoRA: Efficient fine-tuning method that introduces low-rank updates to model weights.
Domain Adaptation	- Unsupervised Domain Adaptation (UDA): Adapts the model to a new domain with limited labeled data.
Handling Imbalanced Data	- SMOTE (Synthetic Minority Over-sampling Technique): Creates synthetic data to balance class distribution.
Handling Long Contexts	- Hierarchical Attention Mechanisms: Process information at different granularities for long sequences.
Ensemble Methods	- Conditional Weighting: Assigns weights to models based on task or input characteristics.
Contrastive Learning	- SimCLR (SimContrastive Learning): Improves representation learning with unlabeled data. - MoELoRA: Combines Mixture of Experts with Low-Rank Adaptation for efficient and robust model fine-tuning.
Knowledge Distillation	- Knowledge Distillation with Attention Transfer: Transfers knowledge from a large teacher to a smaller student model while focusing on important attention patterns.

Hybrid Models	- Neuro-Symbolic Reasoning: Integrates symbolic reasoning with neural networks for improved interpretability and generalization.
Data Augmentation Techniques	- Back-Translation with Noise Injection: Creates more diverse training data by translating text to another language and then back with additional noise.
Active Learning	- Pool-based Active Learning: Selects the most informative data points for human annotation.
Zero-Shot and Few-Shot Learning	- Prompt Tuning with Continuous Hyperparameters: Improves performance on few-shot tasks by tuning prompt parameters.
Evaluation and Validation	- Human-in-the-Loop Evaluation: Integrates human judgment for comprehensive evaluation.
Model Interpretability	- Causal AI Techniques: Explains model decisions by analyzing causal relationships in the data.
Ethical and Legal Considerations	- Algorithmic Bias Detection: Techniques to identify and mitigate biases in AI systems. - Explainability in Regulation: Regulatory requirements for explainability in AI models.
Operational Aspects	- AutoML for Fine-Tuning: Automates hyperparameter tuning and model selection for efficient fine-tuning.

Cost Management	- Cost-Aware Training: Optimizes training for resource efficiency while maintaining performance.
Regularization Techniques	- Spectral Normalization: Improves training stability and generalization.
Data Cleaning and Preprocessing	- Noise Reduction - Text Normalization - Entity Recognition - Contextual Named Entity Recognition (NER): Improves NER accuracy by considering context.
Cross-Lingual and Multilingual Models	- Cross-Lingual Transfer Learning - Multilingual Embeddings - Multilingual Transformers with Knowledge Distillation: Improves performance of multilingual models by leveraging knowledge distillation.
Personalization and Customization	- User-Specific Fine-Tuning: Adapting models based on user-specific data for personalized experiences. - Contextual Adaptation: Customizing responses based on user context and history. - User-Adaptive Fine-Tuning: Tailors the model to individual user preferences.
Interactive Learning	- RLHF (Reinforcement Learning from User Feedback): Continually improving models based on user interactions and feedback. - Human-in-the-Loop Fine-Tuning: Involving human feedback in the fine-tuning process to correct and guide model outputs. - Interactive Prompt Refinement: Enables users to refine prompts iteratively for improved performance.

Scalability and Performance Optimization	- Mixed Precision Training - Model Pruning and Quantization (e.g., GPTQ) - Distillation and Compression (e.g., 1-bit models) - Hardware-Aware Training: Optimizes training for specific GPU platforms.
Robustness and Security	- Adversarial Training: Training models with adversarial examples to improve robustness against attacks. - Robustness Testing: Systematically testing model robustness against various input perturbations. - Security Measures: Implementing measures to protect models from malicious inputs and data breaches.
Ethical AI Development	- Bias Mitigation: Techniques to detect and mitigate biases in models. - Fairness and Accountability: Ensuring models are fair and transparent in their decision-making processes. - Responsible AI Practices: Adhering to ethical guidelines and frameworks for AI development.
Visualization and Analysis Tools	- TensorBoard - Embedding Projector - Explainability Tools (LIME, SHAP), MLflows, Wandb
Integration with Existing Systems	- API Deployment: Exposing model functionalities through APIs for easy integration. - Edge Deployment: Techniques for deploying models on edge devices with limited resources. - Cloud Platforms: Leveraging cloud services for scalable model deployment and management.

Benchmarking and Standardization	- Benchmark Datasets: Using established benchmarks like GLUE, SuperGLUE, and SQuAD to evaluate model performance. - Performance Metrics: Standardizing on key metrics such as perplexity, accuracy, F1 score, and BLEU, BERTScore for consistent evaluation.
Cross-Model Interoperability	- ONNX (Open Neural Network Exchange): Using the ONNX format for model interoperability across different frameworks. - TF Hub - PyTorch Hub

To download complete chart ⬇

https://github.com/Abonia1/Fine-Tuning-LLMs-Key-Concepts-and-Terms