To make sure you're fully prepared to dive into the syllabus for building cutting-edge LLMs and AI agents, here's a focused and prioritized checklist of concepts from your Udacity deep learning course to revisit:

**✅ 1. Neural Network Fundamentals:**

* **Matrix Multiplication & Vectorization**
  + Efficient implementations for model training.
  + Broadcasting in NumPy/PyTorch.
* **Activation Functions & Layers**
  + Common activations: ReLU, sigmoid, tanh, softmax, GELU (modern activations).
  + Understanding layer stacking, hidden layers, input/output shapes.
* **Backpropagation & Gradient Descent**
  + Calculating and interpreting gradients.
  + Stochastic Gradient Descent (SGD), Adam optimizer.

**✅ 2. Model Training and Optimization:**

* **Loss Functions**
  + Cross-entropy, mean-squared error (MSE), and log loss.
* **Optimization Techniques**
  + Regularization (L1/L2), dropout, weight decay.
  + Batch normalization.
  + Hyperparameter tuning (learning rate, batch size).
* **Debugging and Monitoring**
  + Using learning curves, accuracy curves, and loss curves effectively.

**✅ 3. Embeddings & Representation Learning:**

* **Word Embeddings (Word2Vec, GloVe, FastText)**
  + Understanding semantic vector representations.
  + Embedding layers in neural networks.
* **Transfer Learning & Pretraining**
  + Leveraging pretrained embeddings and models.
  + Fine-tuning pretrained models for downstream tasks.

**✅ 4. Sequence Modeling & NLP Basics:**

* **Sequence Modeling Architectures**
  + Recurrent Neural Networks (RNN), LSTMs, GRUs: key strengths and limitations.
  + Managing variable-length sequences (padding, masking).
* **Text Preprocessing & Tokenization**
  + Tokenization methods (subword, wordpiece, byte-pair encoding).
  + Handling special tokens ([PAD], [MASK], [CLS], [SEP], etc.).

**✅ 5. Attention Mechanisms & Transformers:**

* **Attention Mechanism Basics**
  + Concept and purpose of attention.
  + Scaled Dot-Product Attention (query-key-value framework).
* **Transformers (Critical for LLMs)**
  + Architecture overview (multi-head attention, encoder-decoder, positional embeddings).
  + Key papers (especially "Attention Is All You Need").

**✅ 6. Generative Models & Language Modeling Concepts:**

* **Language Modeling Basics**
  + Next-token prediction and autoregressive modeling.
  + Evaluating language models (perplexity).
* **Generative vs. Discriminative Models**
  + Understanding how generative models create text.

**✅ 7. Model Deployment & Practical Considerations:**

* **Model Serving Basics**
  + Packaging models with frameworks (e.g., PyTorch, TensorFlow, HuggingFace).
  + Basic API creation for model endpoints.
* **Computational Efficiency & Scaling**
  + GPU acceleration and parallel processing concepts.
  + Basics of distributed training/inference (conceptual understanding).