Mathematics & Machine Learning Basics

- Linear Algebra (Vectors, Matrices)
- Probability & Statistics (Bayes' Theorem, Gaussian Distribution)
- Optimization (Gradient Descent, Backpropagation)
- Basic ML Algorithms (Logistic Regression, Decision Trees, Random Forest)

Resources

- "Mathematics for Machine Learning" by Deisenroth, Faisal, Ong
- Andrew Ng's ML Course (Coursera)
- Kaggle ML tutorials

NLP Basics

- Tokenization (Word, Subword, Character-Level)
- Lemmatization & Stemming
- Stopwords & Part-of-Speech (POS) Tagging
- Named Entity Recognition (NER)
- Bag of Words (BoW), TF-IDF

Word Embeddings & Feature Engineering

- Word2Vec (CBOW, Skip-Gram)
- GloVe
- FastText

Traditional NLP Models

- Naïve Bayes for Text Classification
- Logistic Regression for Sentiment Analysis
- Latent Dirichlet Allocation (LDA) for Topic Modeling

Resources

- "Speech and Language Processing" by Jurafsky & Martin
- "Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron
- NLP with Python (NLTK, SpaCy, Scikit-learn)

Practice

- Build a spam classifier using Naïve Bayes
- Sentiment analysis on Twitter dataset
- Named Entity Recognition using SpaCy

Neural Networks for NLP

- RNNs, LSTMs, GRUs

- Attention Mechanism
- Sequence-to-Sequence Models

Transformers (Modern NLP Models)

- BERT, RoBERTa, T5
- GPT (GPT-3, GPT-4, LLaMA, Mistral)
- Fine-Tuning Pretrained Language Models

Advanced NLP Techniques

- Text Summarization (Abstractive & Extractive)
- Machine Translation
- Speech-to-Text (Whisper, DeepSpeech)

Resources

- “Natural Language Processing with Transformers” by Lewis Tunstall, Leandro von Werra
- Hugging Face Course (Transformers, Tokenizers)
- Deep Learning Specialization by Andrew Ng

Practice

- Fine-tune BERT for text classification
- Build a chatbot using GPT models
- Create a text summarizer using T5

Building NLP-Powered Applications

- Chatbots (Rasa, OpenAI API)
- AI Writing Assistants (Text Generation)
- AI-Powered Search Engines (ElasticSearch, BM25)

MLOps & Deployment

- Model Serving (FastAPI, Flask)
- Deploying NLP models on AWS/GCP
- Model Monitoring & Scaling

Resources

- Hugging Face Model Hub
- TensorFlow Extended (TFX) & MLflow
- Practical NLP Projects on GitHub

Final Projects

- End-to-End Sentiment Analysis API
- AI-Powered Resume Screener
- Custom Chatbot (Fine-Tuned GPT/BERT)

⏳ Time Commitment Summary

6 Months (Fast Track)

- Daily: 3-4 hours
- Projects: Work on NLP-based apps after 3 months

12 Months (Balanced Pace)

- Daily: 1-2 hours
- Projects: Start small, then scale to bigger projects

Since you want to go deep into NLP, focusing on LLMs, fine-tuning, and AI-driven applications will make you an expert in NLP within a year.

Would you like me to suggest specific project ideas to work on while learning?