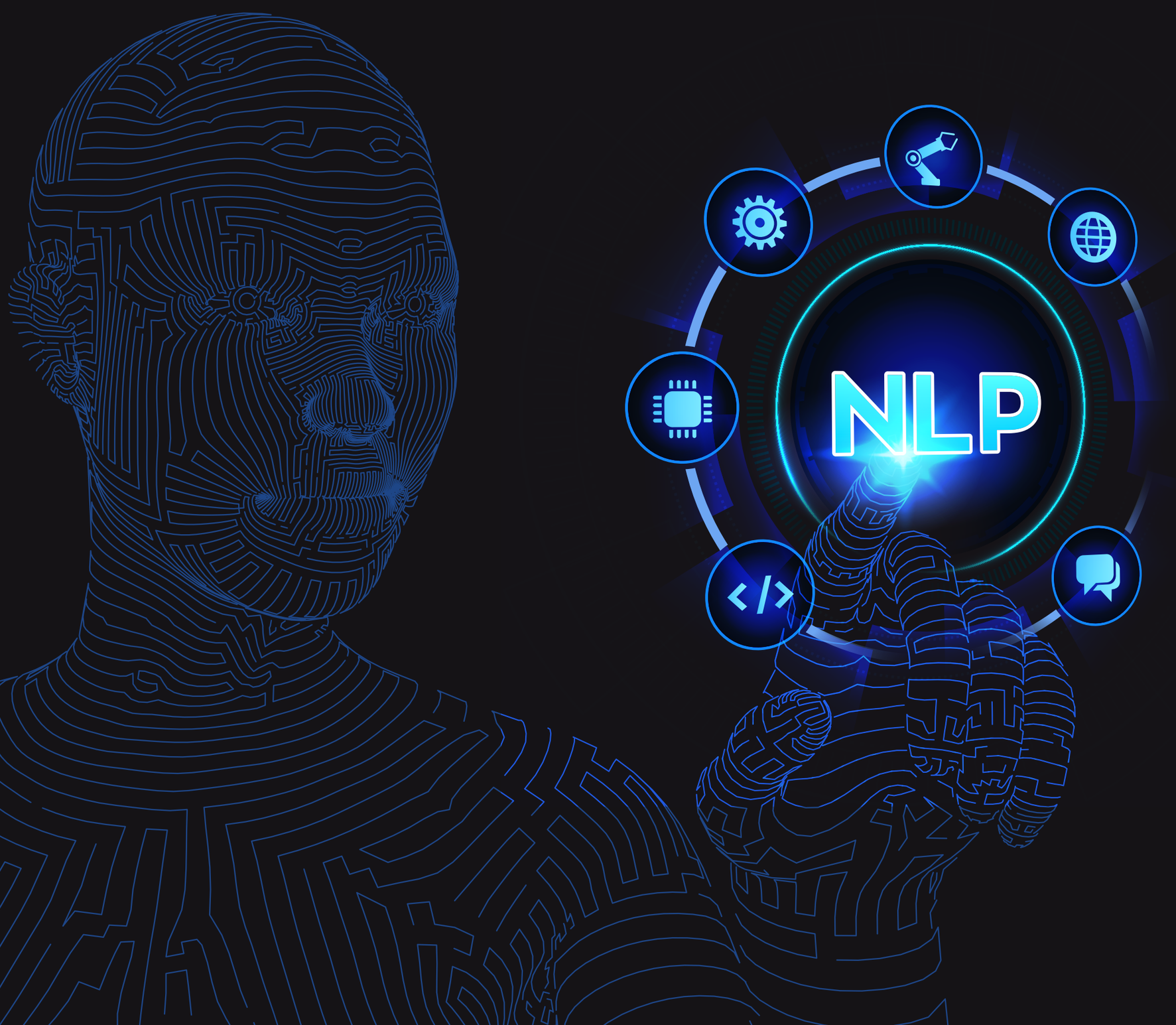
Analytics Vidhya

MASTER NATURAL LANGUAGE PROCESSING

2024



Analytics Vidhya

Learning NLP in 2024 offers the chance to be at the forefront of this technological revolution. It equips you with the tools to train, fine-tune, and implement LLMs, ensuring you're prepared for a future where these models will likely dominate the tech landscape. Therefore, diving into NLP is both a strategic career move and a thrilling intellectual journey, allowing you to contribute to and shape the future of how humans and machines communicate and collaborate.

Analytics Vidhya

PRE-REQUISITES:

- **Python:** Gain proficiency in Python, the primary programming language used in NLP.

- **Basic ML Algorithms:**
  - Linear Regression
  - Logistic Regression
  - KNN
  - Decision Tree
  - Random Forest
  - Naive Bayes
  - Support Vector Machine

- **Basic Deep Learning Concepts:** Learn the fundamentals of neural networks.

- **Mathematics:** Solidify your understanding of statistics and probability.

Analytics Vidhya

MONTH-1

Getting Started with Textual Data

Text Preprocessing:

- Tokenization
- Text Cleaning
- Stopword removal
- Stemming and Lemmatization

Word Embeddings: Understand word representation techniques like

- Word2Vec
- TF-IDF
- One-Hot Encoding

Projects:

- Sentiment Analysis
- Fake News Detection

Research Papers: Read foundational papers on TF-IDF and Word2Vec.

Analytics Vidhya

MONTH-2

Traditional NLP Approaches

Deep Learning NLP Frameworks: Get to grips with PyTorch and TensorFlow.

NLP Concepts: Dive into

- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)
- Long Short-Term Memory Networks (LSTMs)
- Gated Recurrent Units (GRUs)
- Encoder-Decoder Architecture

Projects: Start with text summarization and machine translation projects.

Research Papers: Study CNNs, RNNs, LSTM, GRU, and Encoder-Decoder architecture.

Analytics Vidhya

MONTH-3

Entering into the NEW ERA of NLP

Attention Mechanisms: Study the "Attention Is All You Need" paper and understand the Transformer model.

Familiarize yourself with transformer-based models like BERT, Roberta, Distill Bert, T5, and so on..

Transfer Learning: Learn how to apply pre-trained models to new problems like sentiment analysis, machine translation, summarization, etc.

Projects:

- Next Word Prediction

Research Papers: Read about the “Attention Is All You Need” research paper.

Analytics Vidhya

MONTH-4

Getting Started with LLMs

LLMs Introduction: Learn about different types of Large Language Models (LLMs) and their uses.

Foundation Models: Get acquainted with GPT - 3, Llama 2, Falcon 180B, Mistral 7B and so on.

Prompt Engineering: Understand how to prompt LLMs to get desired outputs effectively.

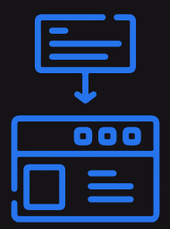
Learn about RAG(Retrieval Augmented Generation system)

Projects: Building LLM applications using RAG

Analytics Vidhya

MONTH-5

Finetuning LLMs



Fine-tuning LLMs: Learn about Fine-tuning foundation models like Llama 2

Learn about Fine-tuning techniques

- Adapter based learning
  - PEFT
  - Lora - Qlora
- Full Model fine-tuning
- Prompt based Fine-tuning
  - One Shot Learning
  - Few Shot Learning

Projects:

- Fine-tune a model for a specific NLP task.

Analytics Vidhya

MONTH-6

Building LLMs from Scratch

Building LLMs: Learn about building LLMs from scratch, considering the latest models and techniques.

Projects: Construct your own LLM, inspired by models like Llama 2.