DevOpsSchool

Course curriculum · November 2026

LEVEL 04 OF 04 · ADVANCED

Language Processing Advanced Training

Build production LLM and retrieval-augmented applications with evaluation, safety controls, efficient inference and monitoring.

LIVE ONLINE **Max 8-10 hours**	CORPORATE (DAY-WISE) **5 days · Days 1-5**	SELF-PACED VIDEO **1-year LMS access**
BATCH SIZE **Max 10 learners**	NEXT BATCHES **1st of every month · weekends: first Saturday**	CERTIFICATE **Verifiable certificate of completion**

Foundation	Essential	Intermediate	Advanced
1 day · Included	2 days · Included	3 days · Included	5 days · This curriculum

INSIDE THIS CURRICULUM

What you will achieve	Outcomes and level focus
Who should attend	Audience and prerequisites
Agenda	46 topics · Foundation to Advanced
Hands-on labs	4 labs
Formats and fees	Live, 1:1, self-paced, corporate
Trainer and reviews	Rajesh Kumar
Enrolment	Batches, payment, contacts

WHY DEVOPSSCHOOL

- Rated 4.6 on Google
- 1,500+ verifiable certificates issued in the public registry
- Trainer Rajesh Kumar: 20 years in industry
- Small batches: max 10 learners, every session recorded

Course page: www.devopsschool.com/courses/languageprocessing/language-processing-advance-training.html
Talk to us: WhatsApp +91 99057 40781 · +1 (469) 756-6329 · contact@DevOpsSchool.com · reply within 1 hour

OUTCOMES

What you will achieve

By the end of the Language Processing course path, you will be able to:

1

Process and represent text

Clean, tokenize and vectorise real text data with Python NLP libraries.

2

Build and evaluate models

Train, fine-tune and measure classic and transformer NLP models for classification, extraction and search.

3

Deliver production NLP

Deploy, evaluate and monitor LLM and retrieval-augmented applications responsibly.

WHAT THE ADVANCED LEVEL ADDS

LLMs, RAG, evaluation and production NLP

Hands-on: Ship a RAG question answering service. Index a document collection in a vector store, build a retrieval-augmented answering API with citations, evaluate it against a question set and add latency and drift monitoring.

WHO SHOULD ATTEND

- Experienced Language Processing practitioners responsible for quality, security and scale.
- Senior engineers, architects and leads who design and troubleshoot Language Processing solutions.
- Teams making Language Processing an organisation-wide standard.

BEFORE YOU START

- The Intermediate level, or equivalent hands-on experience.
- You can already: Use word embeddings and transformer models, fine-tune them on your own data and serve predictions through an API.
- Basic computer skills and a laptop with a stable internet connection for the live sessions and labs.
- This curriculum includes every earlier level, so it is complete on its own.

AGENDA · FOUNDATION TO ADVANCED

Language Processing Advanced Training agenda

46 topics. Each level builds on the one before it, so this curriculum lists every earlier level in full. Days are for corporate (day-wise) delivery; live online batches cover the same agenda in max 8-10 hours.

DAY 1 — Foundation: NLP concepts, Python setup and text preprocessing

Included from Foundation

1. What natural language processing is: tasks such as classification, extraction, search, translation and generation
2. NLP pipeline overview: data collection, preprocessing, representation, modelling, evaluation and deployment
3. Environment setup: Python, virtual environments, Jupyter, NLTK, spaCy, scikit-learn and Hugging Face libraries
4. Working with text in Python: Unicode, encodings, string methods and reading files, CSV and JSON
5. Regular expressions for text: patterns, groups, re.findall and re.sub for cleaning and extraction
6. Tokenization: word, sentence and subword tokens with NLTK, spaCy and Hugging Face tokenizers
7. Normalization: lowercasing, punctuation, stop words, stemming (Porter, Snowball) and lemmatization
8. Linguistic basics: morphology, syntax, part-of-speech tags and dependency structure
9. Corpora and datasets: NLTK corpora, Hugging Face Datasets and building your own labelled corpus
10. Counting and exploring text: frequency distributions, n-grams, collocations and word clouds
11. First working example: a cleaned word-frequency report from a set of documents

DAY 2 — Essential: Text representation, classic models and spaCy

Included from Essential

12. Bag-of-words and TF-IDF: CountVectorizer, TfidfVectorizer, vocabulary size and sparse matrices
13. Text classification: Naive Bayes, logistic regression and linear SVM pipelines in scikit-learn
14. Evaluation: train and test split, cross-validation, precision, recall, F1 and confusion matrices
15. Sentiment analysis: lexicon approaches (VADER) versus supervised models
16. spaCy pipelines: nlp objects, Doc, Token and Span, pipeline components and model packages
17. Syntax analysis: part-of-speech tagging and dependency parsing with spaCy and displaCy visualisation
18. Named entity recognition: pretrained NER, entity labels and rule-based EntityRuler patterns
19. Rule-based matching: spaCy Matcher and PhraseMatcher for domain terms and patterns
20. Text similarity: cosine similarity, document clustering with k-means and topic modelling with LDA
21. Handling imbalanced and noisy text data: resampling, class weights and error analysis
22. Saving and reusing models: joblib pipelines and spaCy model packaging

DAY 3 — Intermediate: Embeddings, transformers and fine-tuning

Included from Intermediate

23. Word embeddings: Word2Vec, GloVe and fastText, training with Gensim and exploring analogies
24. Sequence models: RNN, LSTM and GRU for text, and why attention replaced them
25. Transformer architecture: self-attention, encoder, decoder and encoder-decoder models
26. Hugging Face Transformers: pipelines, AutoTokenizer, AutoModel and the model hub
27. Fine-tuning for classification: BERT-family models with the Trainer API and the evaluate library
28. Token classification: fine-tuning for NER and handling subword label alignment
29. Sentence embeddings and semantic search: sentence-transformers, vector similarity and FAISS indexes
30. Summarisation and translation: sequence-to-sequence models such as T5 and BART

31. Question answering: extractive QA with SQuAD-style models
32. Multilingual NLP: multilingual models, language detection and tokenisation issues
33. Serving NLP models: FastAPI endpoints, batching, model loading and Docker packaging
34. Troubleshooting: tokenizer mismatches, sequence length limits, overfitting and GPU memory errors

DAYS 4-5 **Advanced: LLMs, RAG, evaluation and production NLP** **New at this level**

35. Large language models: decoder-only architecture, context windows, tokens, temperature and sampling
36. Prompt engineering: instructions, few-shot examples, structured JSON output and function or tool calling
37. Retrieval-augmented generation: chunking, embeddings, vector databases, hybrid search and re-ranking
38. Parameter-efficient fine-tuning: LoRA and QLoRA with the PEFT library on open models
39. Evaluating NLP and LLM systems: BLEU, ROUGE, BERTScore, human review and LLM-as-judge with test sets
40. Hallucination, bias and safety: grounding, citations, guardrails, PII redaction and content filtering
41. Efficient inference: quantisation, distillation, ONNX Runtime and serving engines such as vLLM
42. Scaling pipelines: batch processing with spaCy nlp.pipe, multiprocessing and GPU utilisation
43. MLOps for NLP: experiment tracking with MLflow, model registry, versioned datasets and CI for models
44. Monitoring in production: latency, cost per request, data drift and feedback loops
45. Governance: data licensing, privacy, audit trails and responsible AI review
46. Capstone: an end-to-end document question answering service with evaluation and monitoring

PRACTICE

Hands-on labs

You build these yourself during the course, guided by the trainer.

FOUNDATION

Build a text preprocessing pipeline

Write a Python pipeline that loads documents, cleans them with regex, tokenizes, removes stop words and lemmatizes, then reports top terms and bigrams.

ESSENTIAL

Classify support tickets and extract entities

Train a TF-IDF and logistic regression classifier for support tickets, evaluate it with F1 and a confusion matrix, and add spaCy NER to pull product and customer names.

INTERMEDIATE

Fine-tune and serve a transformer classifier

Fine-tune a DistilBERT model on a labelled dataset, compare it with the TF-IDF baseline and serve it behind a FastAPI endpoint in a Docker container.

ADVANCED

Ship a RAG question answering service

Index a document collection in a vector store, build a retrieval-augmented answering API with citations, evaluate it against a question set and add latency and drift monitoring.

HOW YOU CAN LEARN

Formats and fees

The same agenda in four formats. Fees are in Indian rupees as of November 2026; the course page shows them in your currency.

FORMAT	DURATION	WHAT YOU GET	FEE
Live & interactive (online)	Max 8-10 hours	Instructor-led public batch of max 10 learners with demos, labs, Q&A and assignment review. Weekday batches start on the 1st of every month, weekend batches on the first Saturday of every month.	**INR 34,999**
1:1 live training (online)	Max 8-10 hours, one-to-one	The same agenda, taught one-to-one by the trainer. Starts immediately, without waiting for the next batch; sessions planned around your working hours, at your pace.	**INR 99,999**
Self-paced video	1-year LMS access	All four Language Processing levels, plus 20+ more courses and 50+ tools, at your own pace.	**INR 833/month** billed yearly at INR 9,996
Corporate (day-wise)	5 days for this level	A private batch for your team, on demand. Foundation 1 day, Essential 2, Intermediate 3, Advanced 5 (cumulative).	**Quote on request**

ASSESSMENT AND CERTIFICATE

Your progress is checked through hands-on labs and assignments at every level. On completion you receive a DevOpsSchool certificate of completion with a unique credential ID that anyone can verify at devopsschool.com/certificates.

INCLUDED WITH EVERY LIVE BATCH

- Every live session is recorded and stays in your LMS for a year.
- Can't attend your batch? Move to next month's batch.
- Free demo class before you pay.
- Free 10-15 minute call with an advisor or the trainer to check the course fits you.

YOUR TRAINER

Learn from Rajesh Kumar

Principal Architect, trainer and mentor in DevOps, SRE, DevSecOps and platform engineering, with 20 years across 8+ software MNCs, including Staff DevOps/SRE Engineer for a leading Japanese mobile-payment platform and Principal DevOps Architect and Manager in Europe.

Profile: www.rajeshkumar.xyz · Videos: youtube.com/@TheDevOpsSchool

20 years in industry	**200+** corporate trainings
10,000+ engineers mentored	**70+** organisations coached

REVIEWS

What learners say

Rated 4.6 on Google · more reviews at devopsschool.com/testimonials and in our video testimonials.

> “Rajesh is very good trainer. Rajesh was able to resolve our queries and question effectively. We really liked the hands-on examples covered during this training program.”
>
> Indrayani, India

> “The training was very useful and interactive. Rajesh helped develop the confidence of all.”
>
> Abhinav Gupta, Pune

START YOUR LANGUAGE PROCESSING TRAINING

Join the next batch

Weekday batches start on the 1st of every month and weekend batches on the first Saturday of every month, with max 10 learners each. 1:1 training can start immediately; corporate batches run on demand. We reply within 1 hour.

Pay by UPI, credit and debit cards, net banking or PayPal. Before you pay, read our refund policy. All levels of this course: www.devopsschool.com/courses/languageprocessing/

ENROL OR RESERVE A SEAT
Language Processing Advanced Training

WHATSAPP · INDIA
+91 99057 40781

WHATSAPP · USA
+1 (469) 756-6329

EMAIL
contact@DevOpsSchool.com