

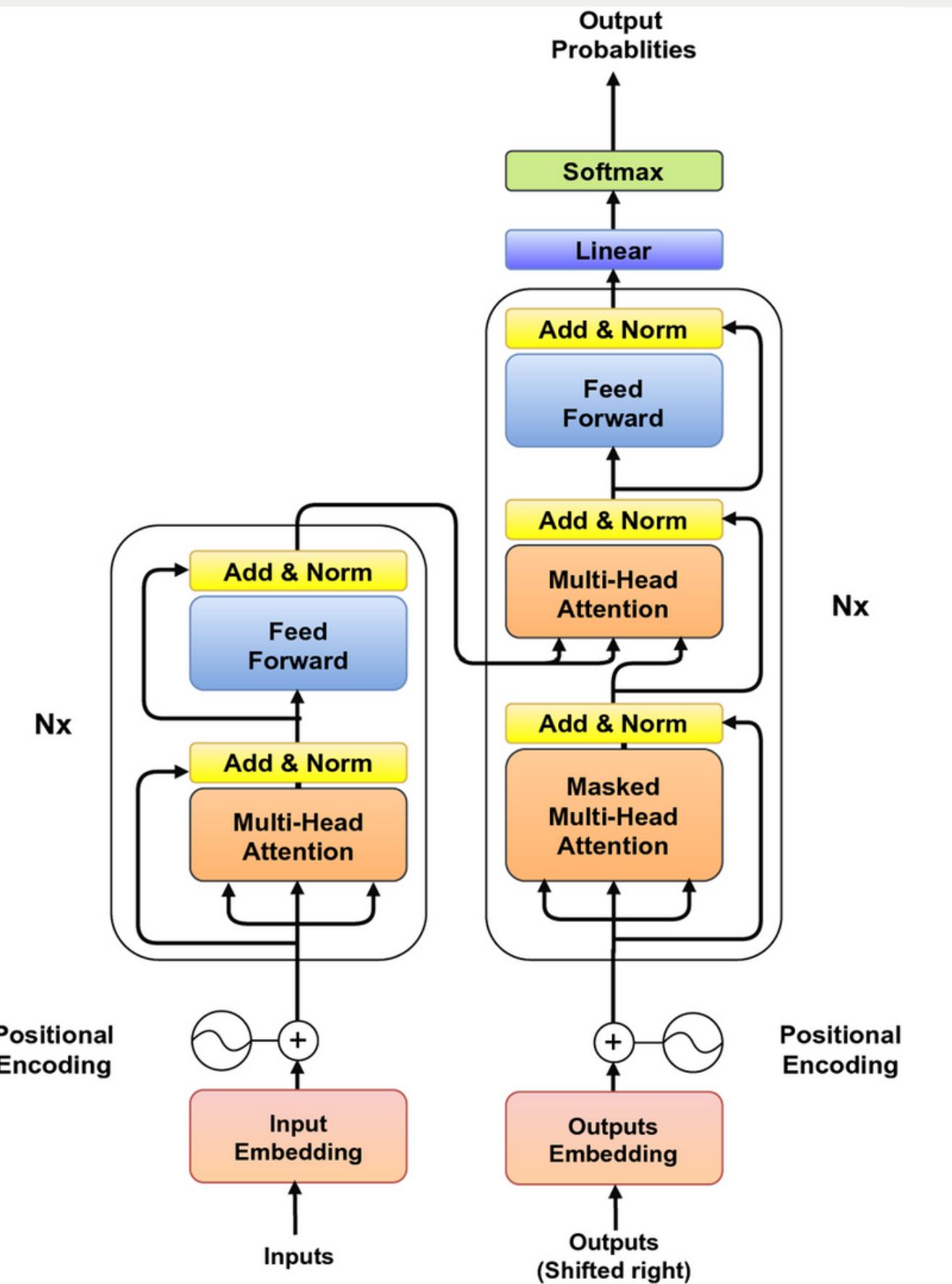


Input Embeddings in Transformers

A Brief Overview

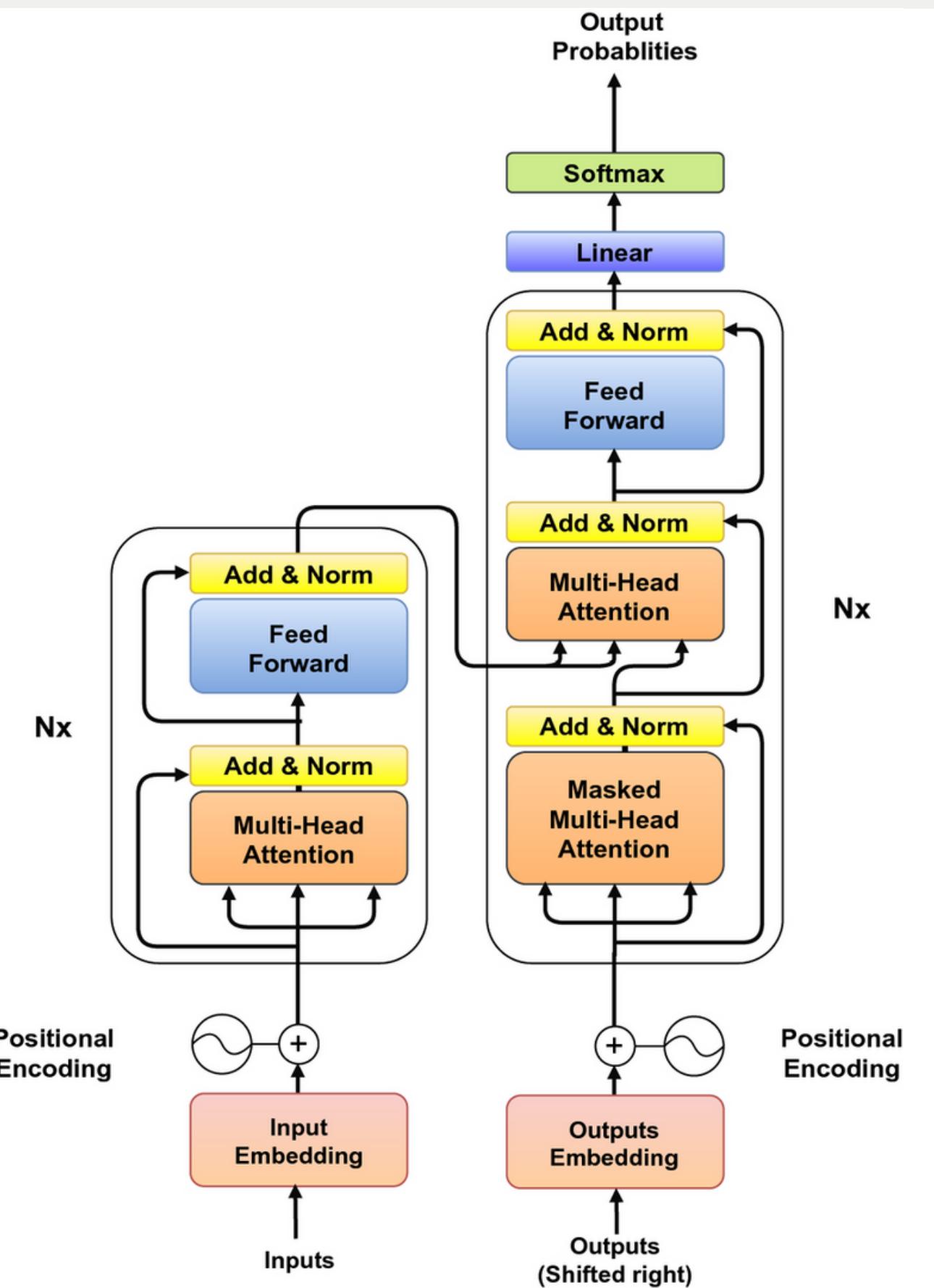
Presented by:
Boniface Samuel Ijeoma

The Transformer Architecture

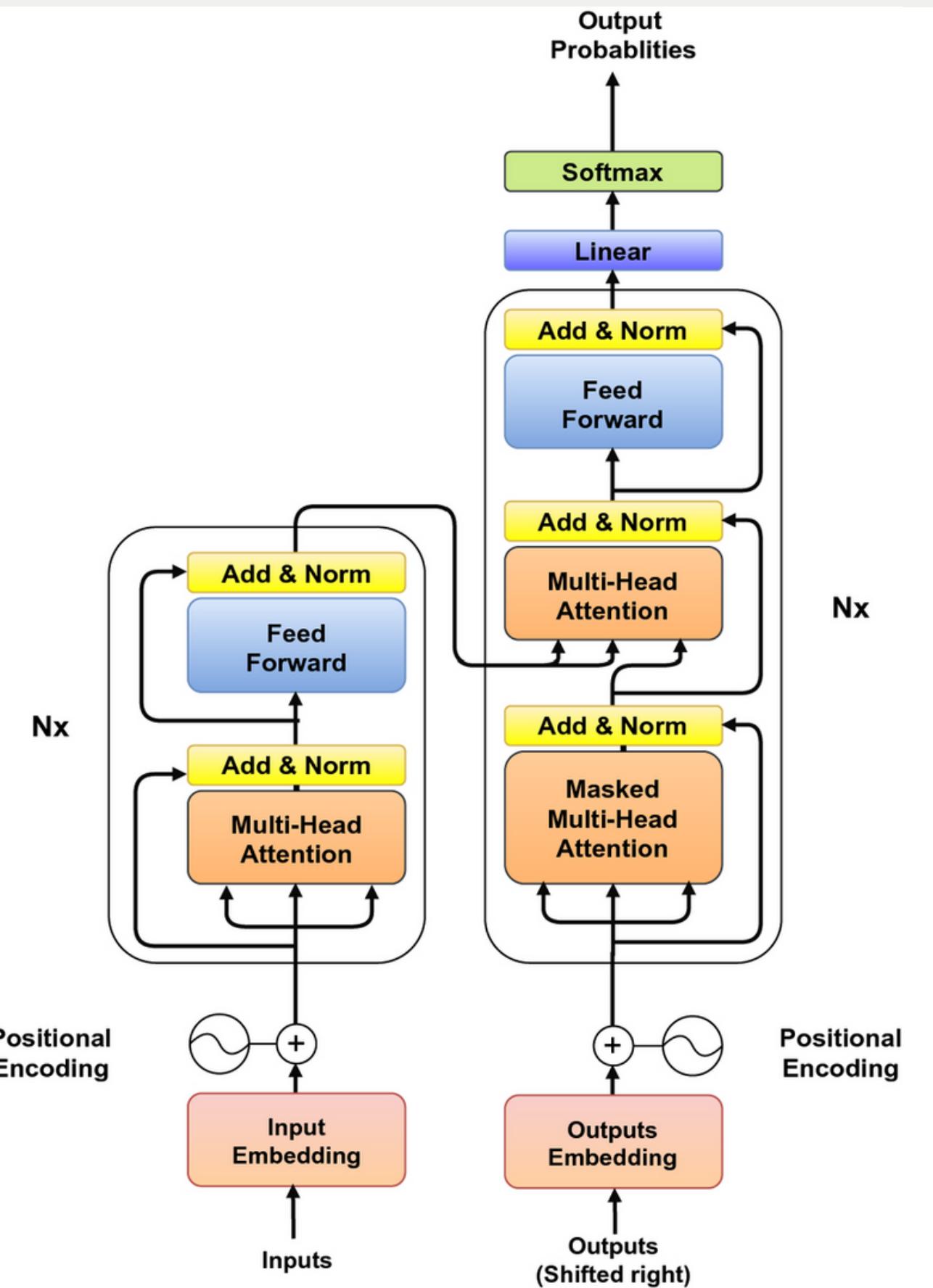


The Transformer Architecture

Introduction



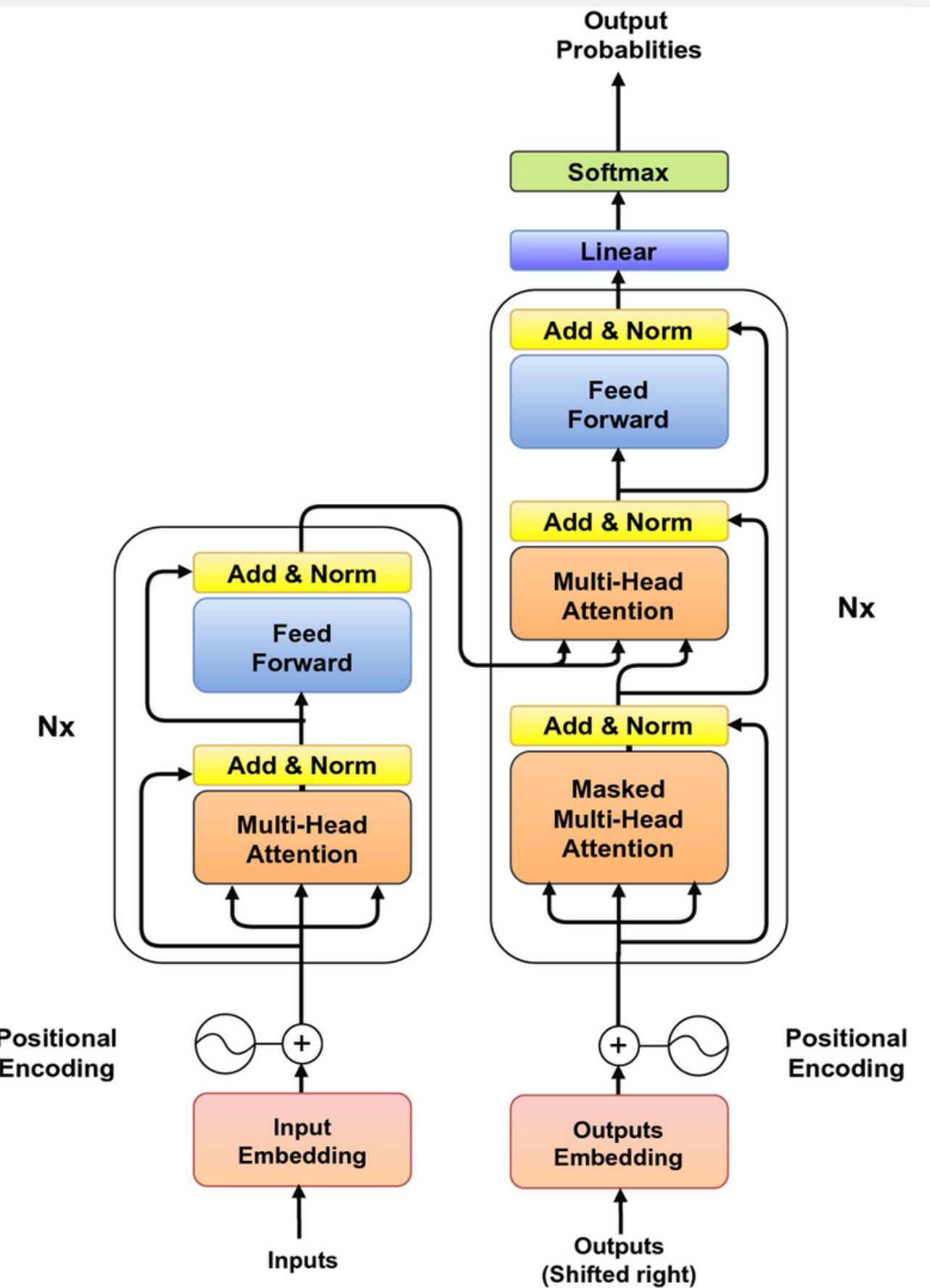
The Transformer Architecture



Introduction

- Introduced 2017 "Attention Is All You Need" by Vaswani

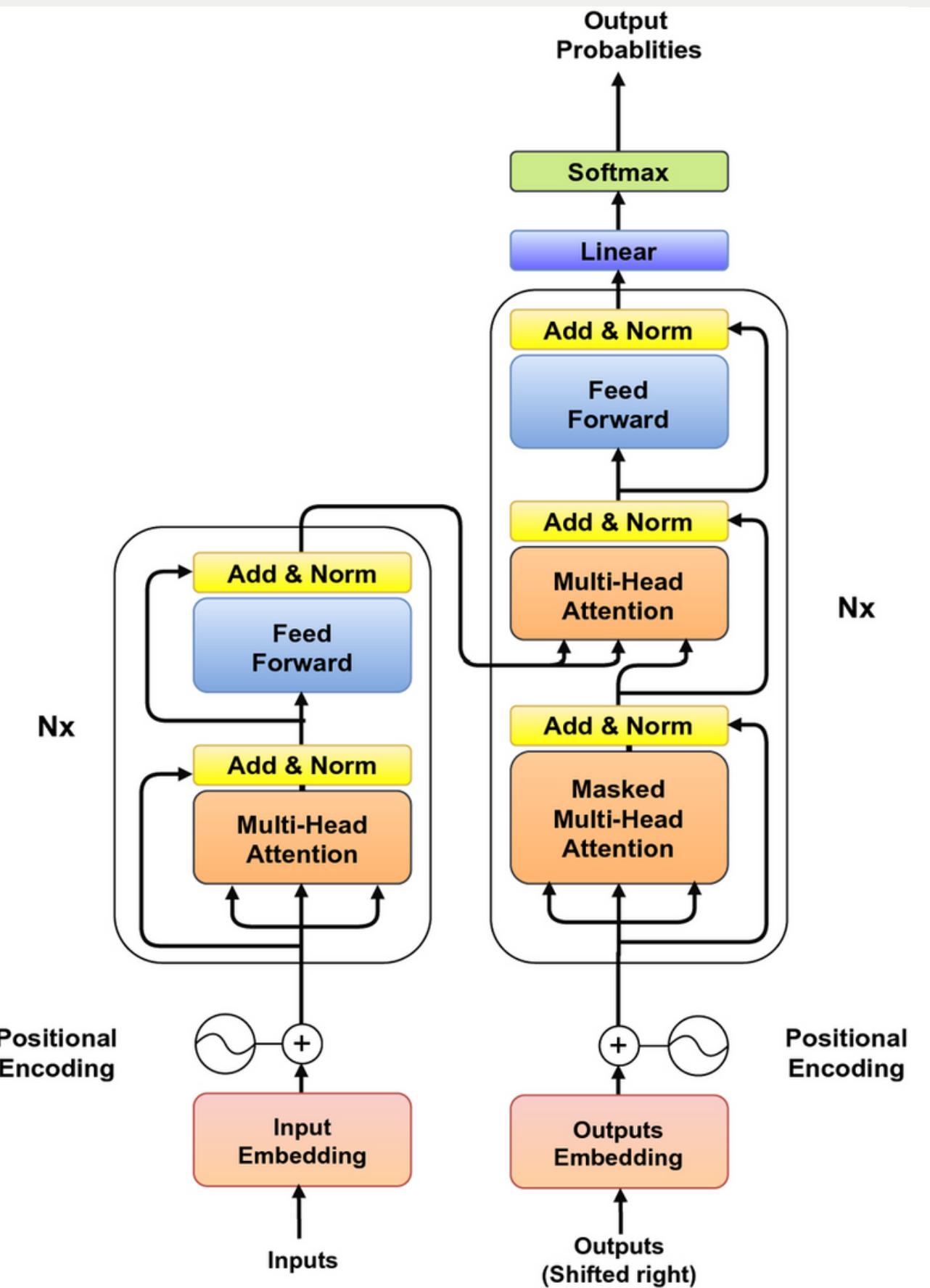
The Transformer Architecture



Introduction

- Introduced 2017 "Attention Is All You Need" by Vaswani
- NLP models primarily utilized RNNs and CNNs for sequential text processing

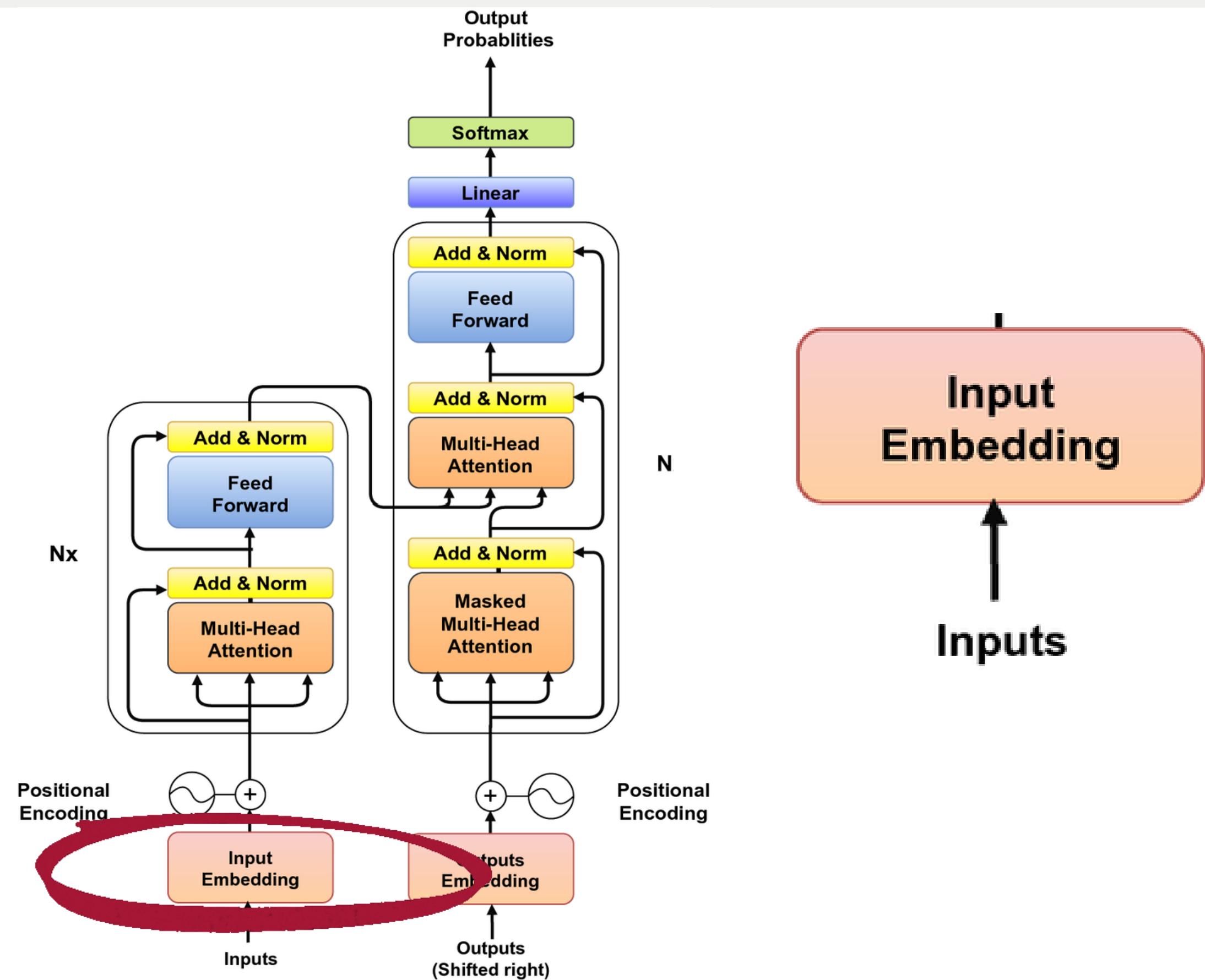
The Transformer Architecture



Introduction

- Introduced 2017 "Attention Is All You Need" by Vaswani
- NLP models primarily utilized RNNs and CNNs for sequential text processing
- Core Innovation: Transformer Architecture

The Transformer Architecture



Introduction

- Introduced 2017 "Attention Is All You Need" by Vaswani
- NLP models primarily utilized RNNs and CNNs for sequential text processing
- Core Innovation: Transformer Architecture

Word Embeddings

Tokenization

Word Embeddings

"This is an input text."

Tokenization

Word Embeddings

Tokenization

- Token
- Token IDs

"This is an input text."



[S]	This	is	an	input	[EOS]
9088	6473	6384	4930	4741	9089

Embeddings Matrix

0	0.123	0.254	0.325	0.985	...	0.214
1	0.257	0.357	0.297	0.998	...	0.558
4741	0.547	0.326	0.258	0.325	...	0.357
4930	0.958	0.355	0.224	0.214	...	0.002
6384	0.247	0.265	0.854	0.245	...	0.124
6473	0.328	0.265	0.658	0.336	...	0.115
9088	0.257	0.368	0.267	0.234	...	0.297
9089	0.231	0.585	0.478	0.345	...	0.327
9999	0.218	0.357	.0159	0.285	...	0.129

Tokenization

- Token
- Token IDs

Word Matrix

Embeddings Matrix

0	0.123	0.254	0.325	0.985	...	0.214
1	0.257	0.357	0.297	0.998	...	0.558
4741	0.547	0.326	0.258	0.325	...	0.357
⋮	⋮	⋮	⋮	⋮	⋮	⋮
4930	0.958	0.355	0.224	0.214	...	0.002
⋮	⋮	⋮	⋮	⋮	⋮	⋮
6384	0.247	0.265	0.854	0.245	...	0.124
⋮	⋮	⋮	⋮	⋮	⋮	⋮
6473	0.328	0.265	0.658	0.336	...	0.115
⋮	⋮	⋮	⋮	⋮	⋮	⋮
9088	0.257	0.368	0.267	0.234	...	0.297
⋮	⋮	⋮	⋮	⋮	⋮	⋮
9089	0.231	0.585	0.478	0.345	...	0.327
⋮	⋮	⋮	⋮	⋮	⋮	⋮
9999	0.218	0.357	.0159	0.285	...	0.129

- Embedding Matrix: vectors that represent words in multi-dimensional space.
- Is an indexed lookup table
- It has a dimension (vocab_size, embedding_dim)

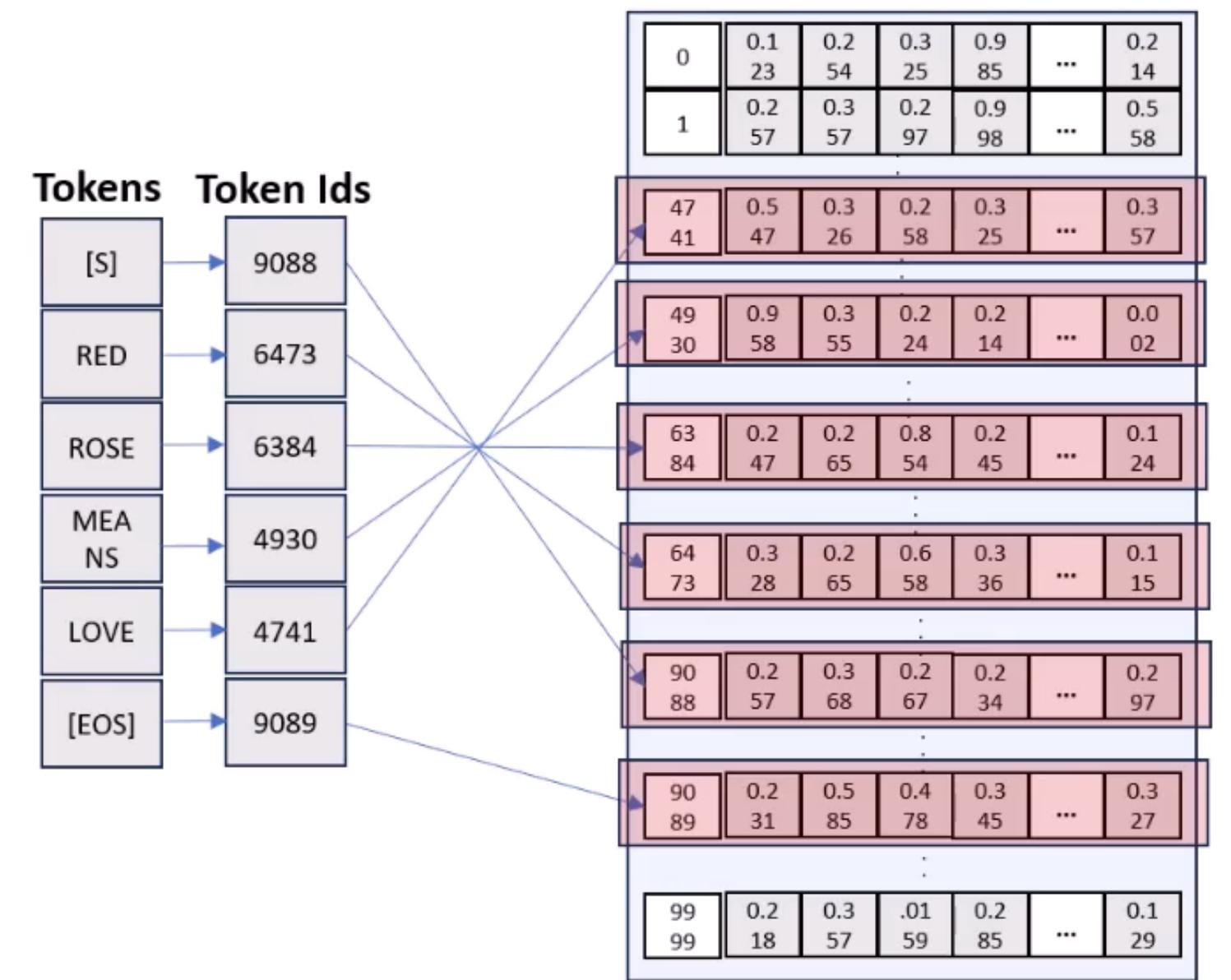
Tokenization

- Token
- Token IDs

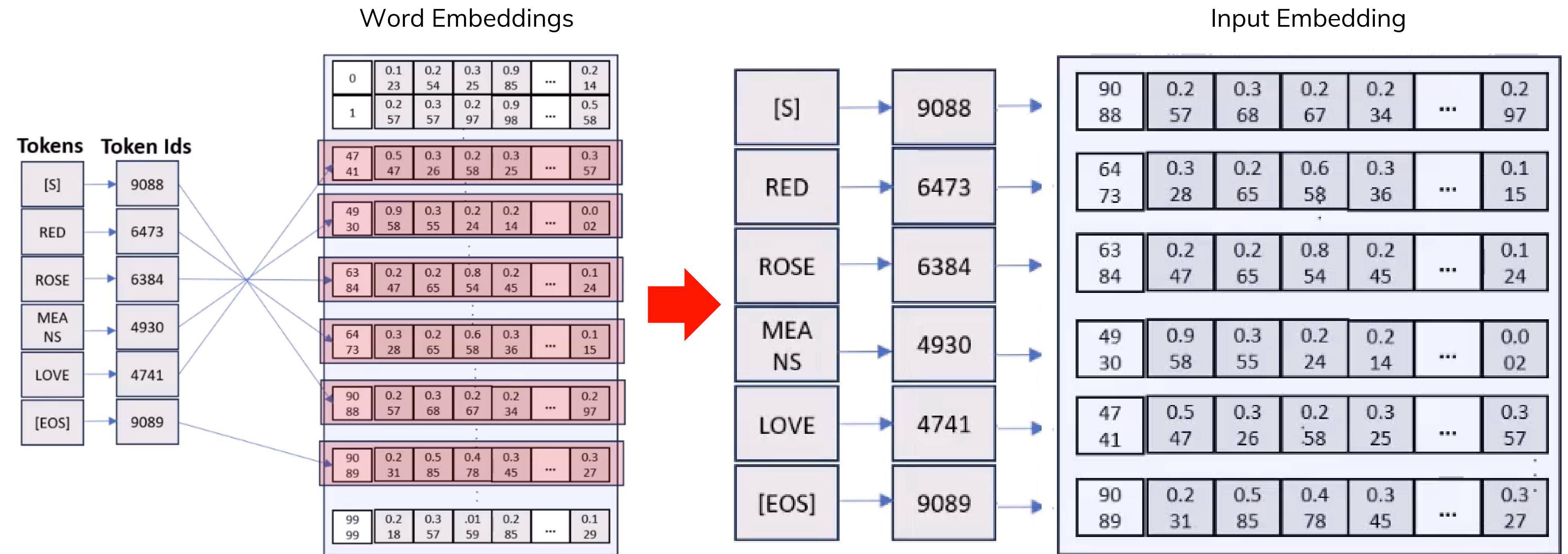
Word Matrix

Input Embeddings

Word Embeddings

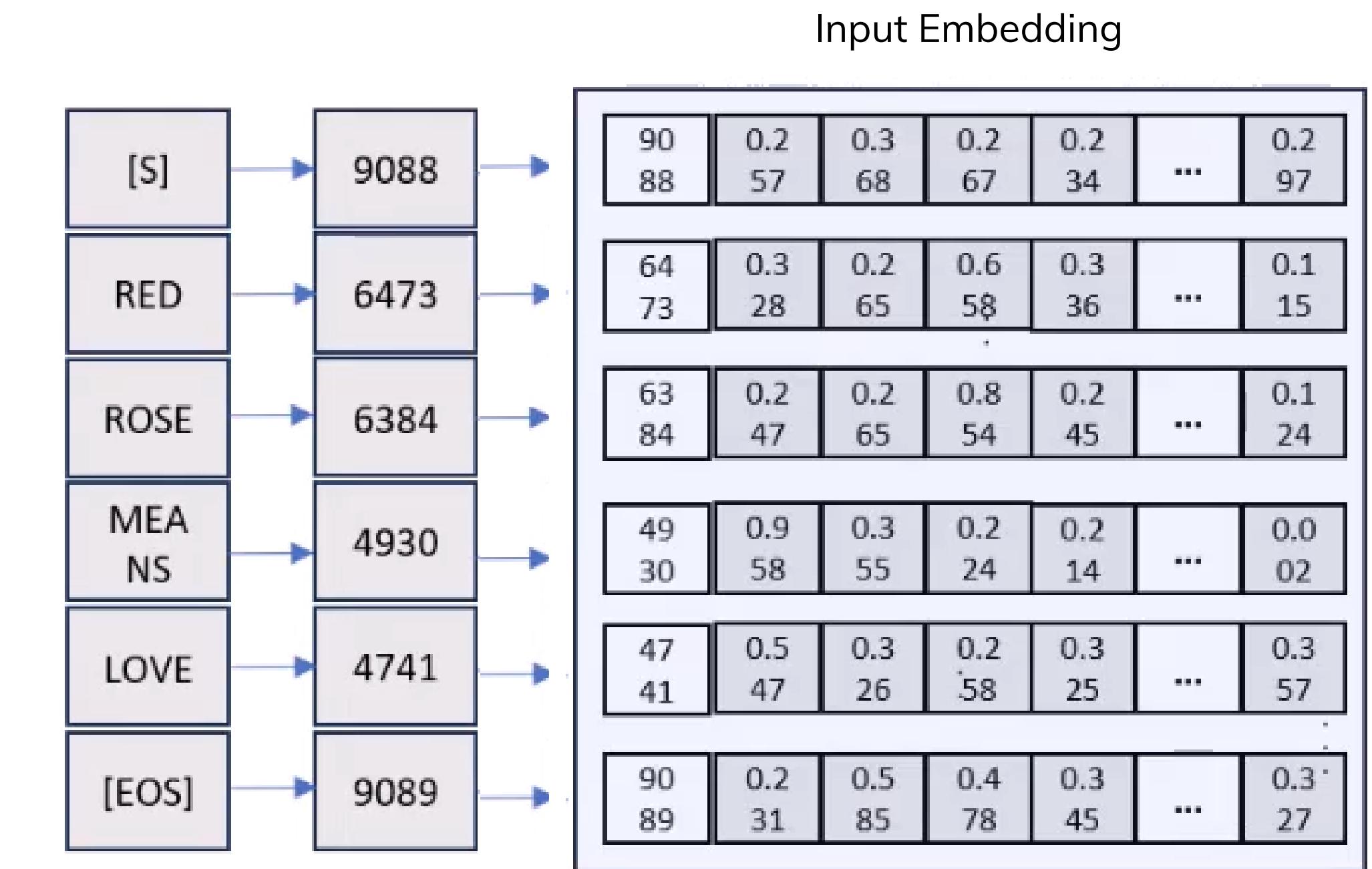


Input Embeddings



Importance

- Efficiency and Dimensionality Reduction
- Enabling Advanced NLP Tasks
- Semantic Representation



Mask-RCNN

YOLOv8

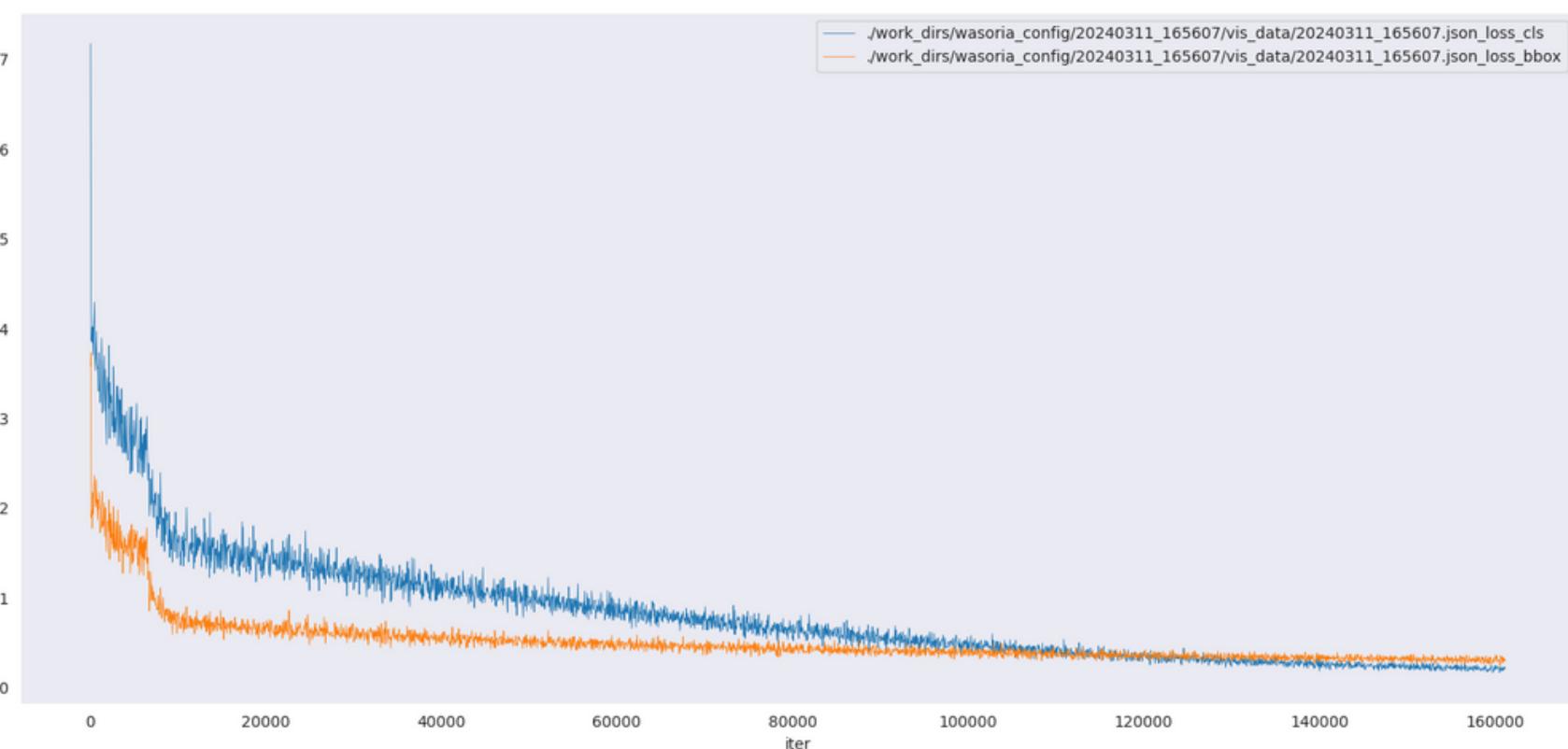
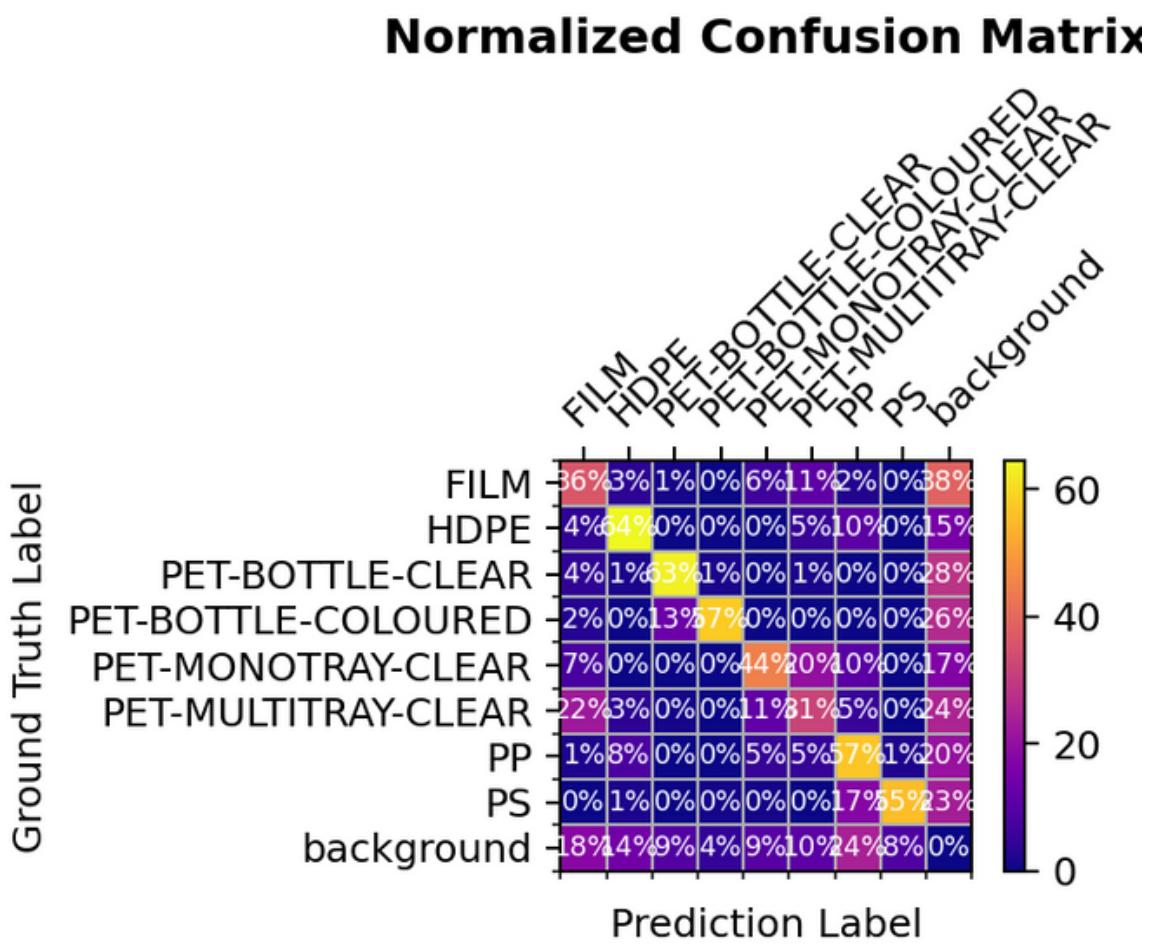
Training Parameters

- 200 Epochs
- 4 Batch Size
- 0.0001 Learning Rate

Dataset

Wasoria_custom_dataset

Mask-RCNN



YOLOv8

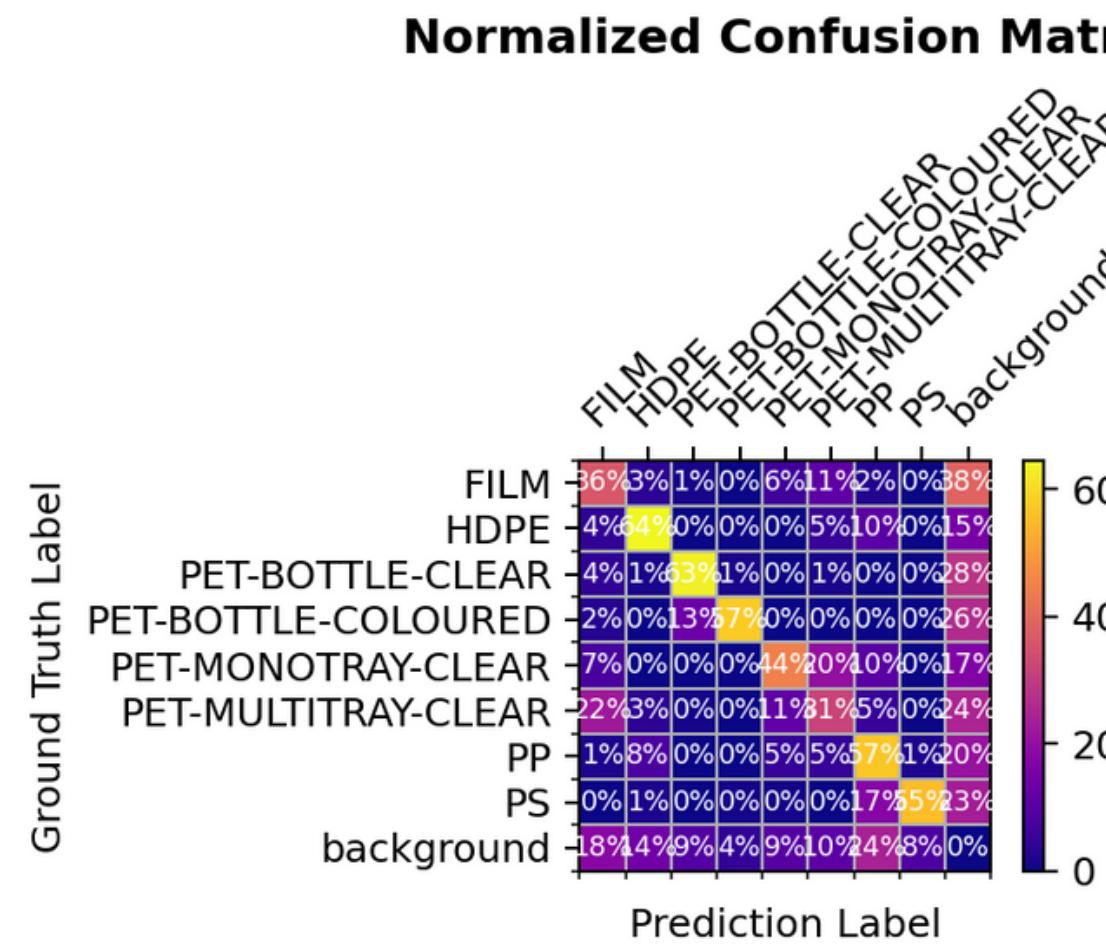
Training Parameters

- 200 Epochs
- 4 Batch Size
- 0.0001 Learning Rate

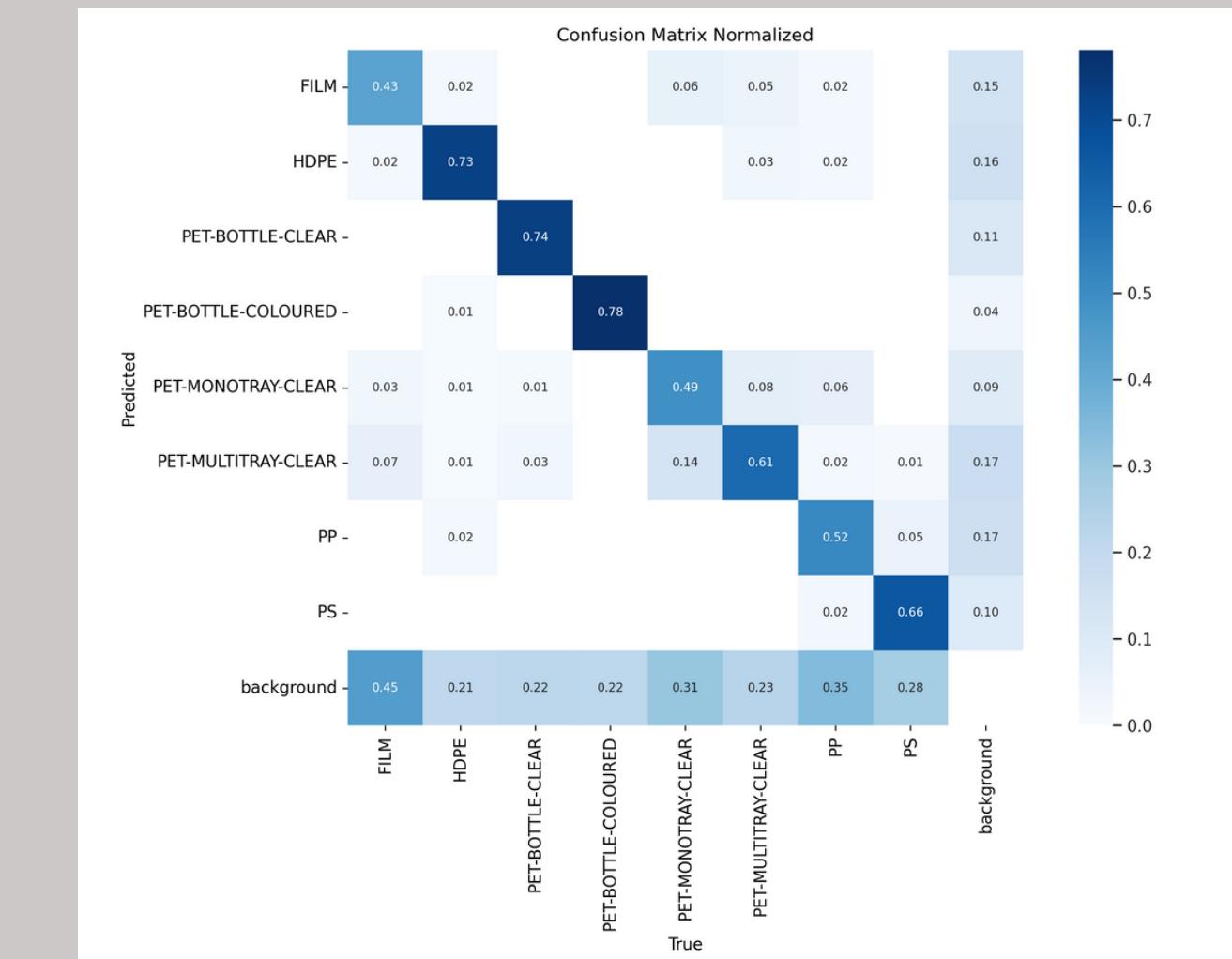
Dataset

Wasoria_custom_dataset

Mask-RCNN



YOLOv8

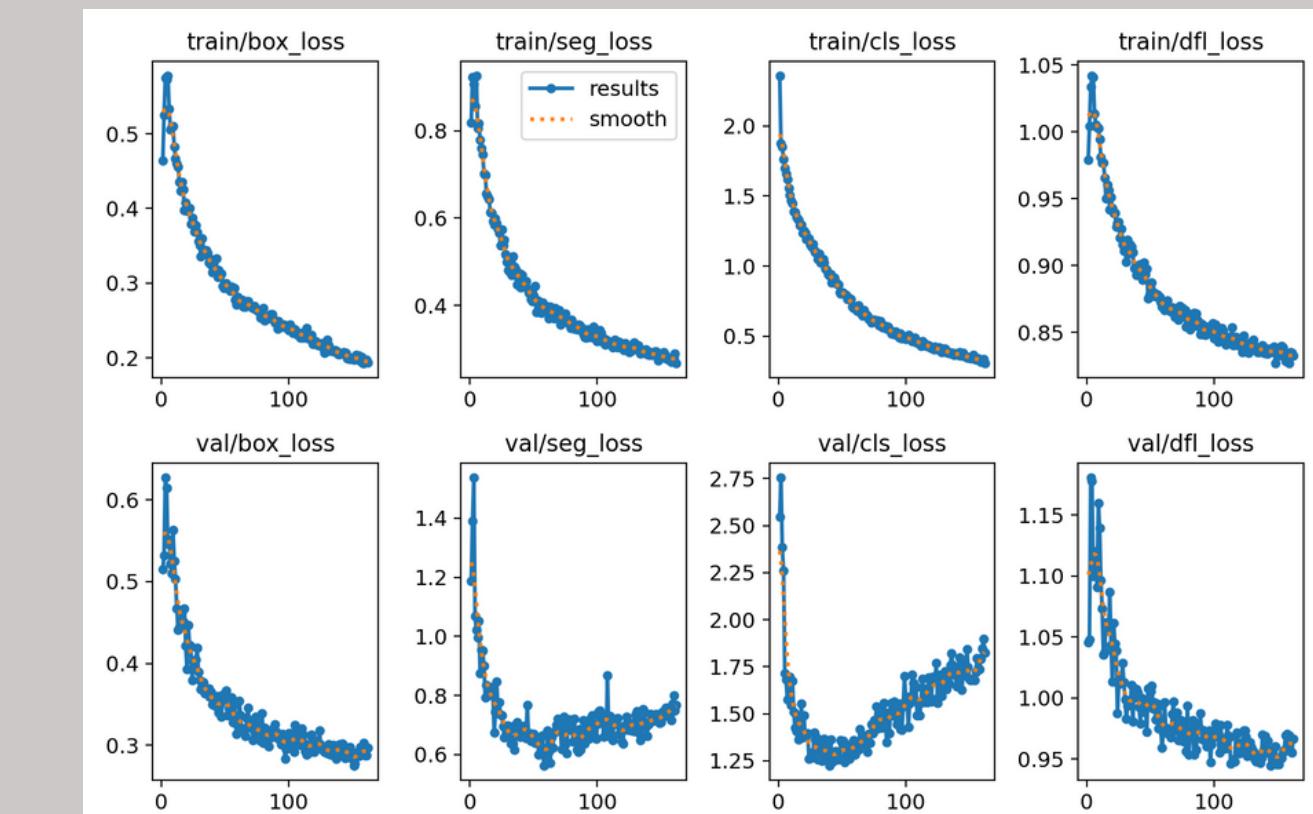
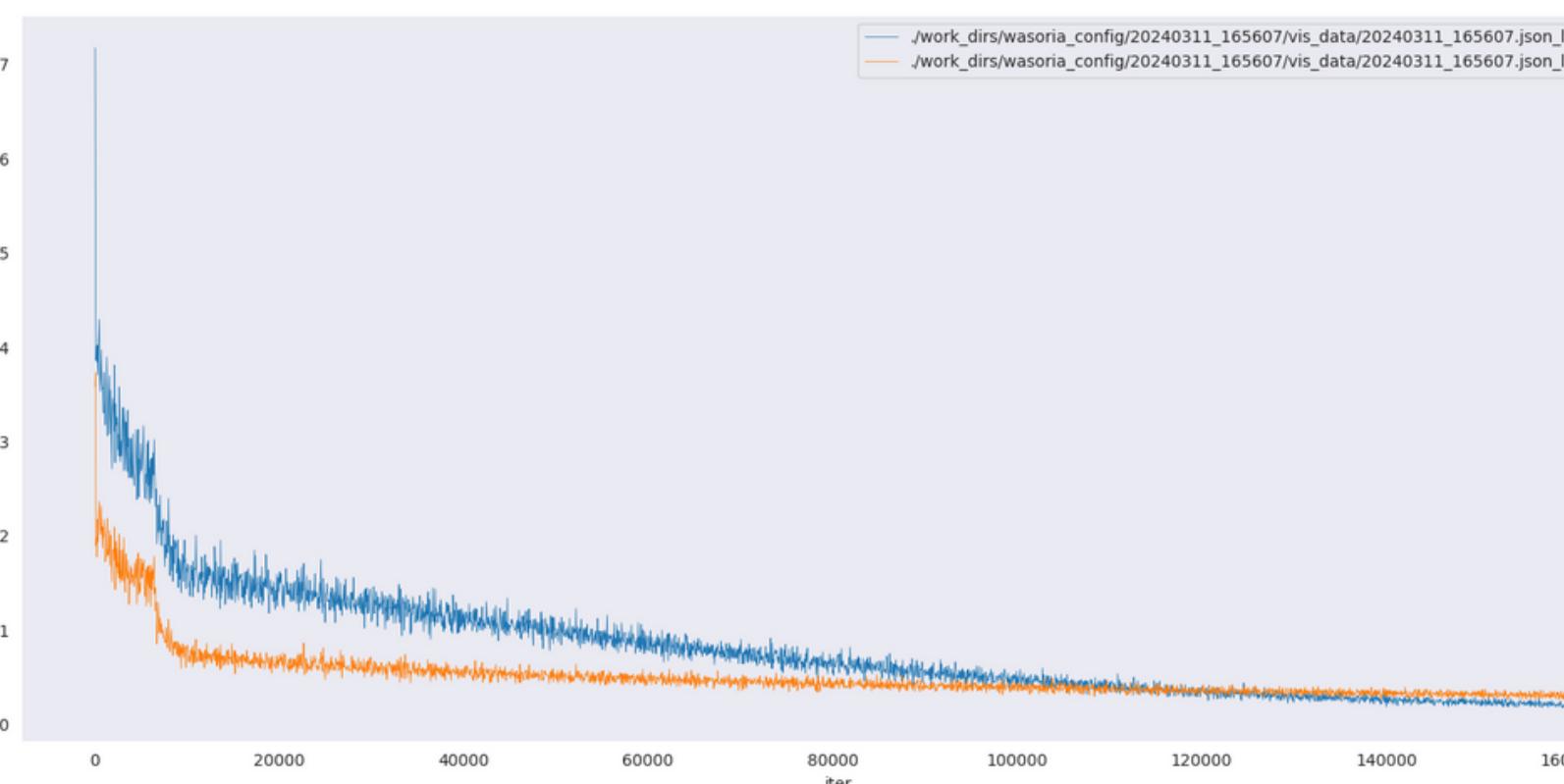


Training Parameters

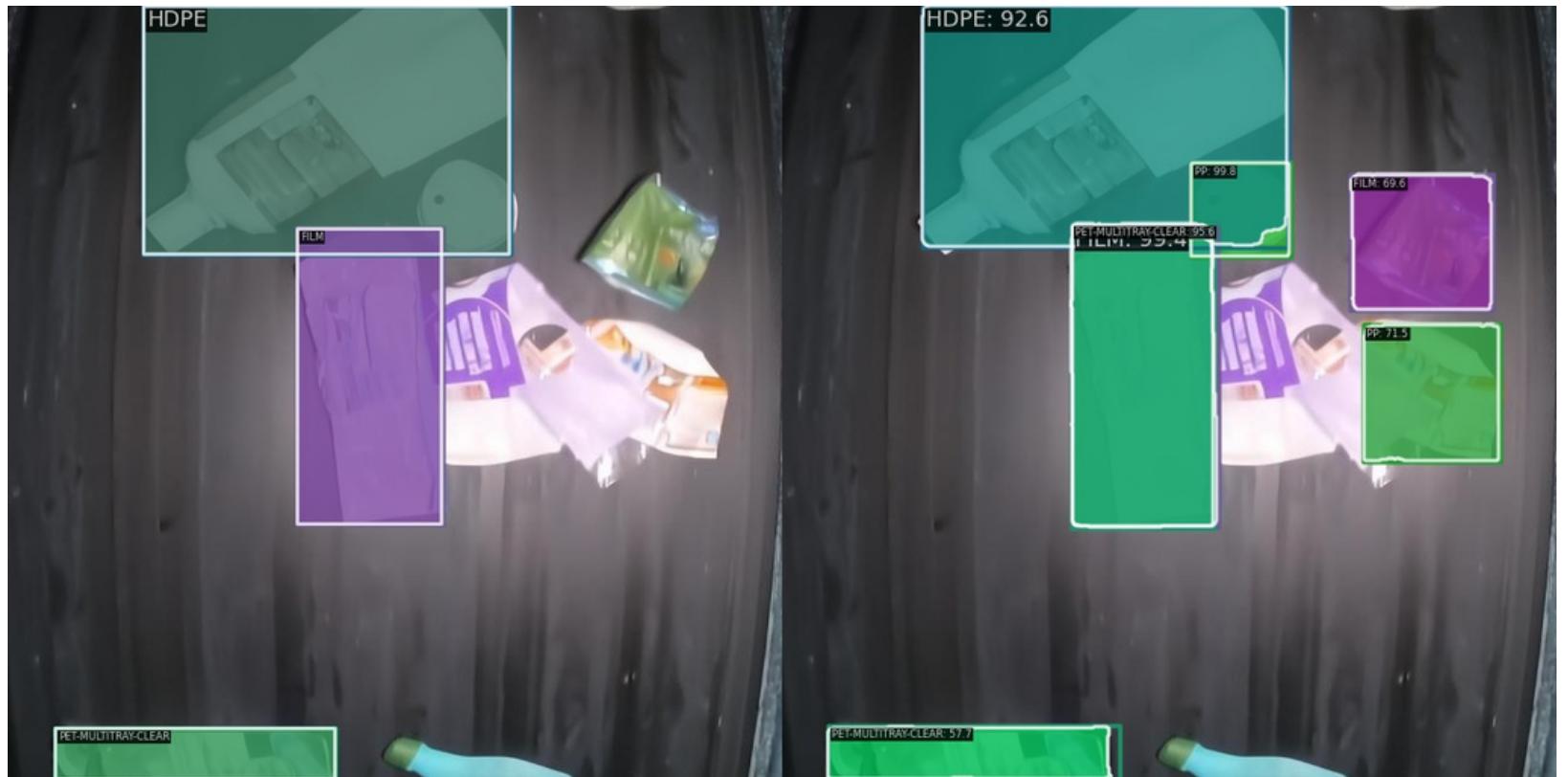
- 200 Epochs
- 4 Batch Size
- 0.0001 Learning Rate

Dataset

Wasoria_custom_dataset



Mask-RCNN



YOLOv8

Training Parameters

- 200 Epochs
- 4 Batch Size
- 0.0001 Learning Rate

Dataset

Wasoria_custom_dataset

Mask-RCNN



YOLOv8



Training Parameters

- 200 Epochs
- 4 Batch Size
- 0.0001 Learning Rate

Dataset

Wasoria_custom_dataset

Reference

The Transformer layer by layer
(Juan Olano)

MMDETECTION (OpenMMLab)
https://mmdetection.readthedocs.io/en/latest/user_guides/deploy.html

Ultralytics
<https://github.com/ultralytics/ultralytics>

THANK YOU
FOR YOUR
ATTENTION

