# English Grammar Correction

## ENCODER-DECODER ARCHITECTURES


SARTHAK DUMBARE (202201040038)
DARSHAN RAMAGADE (202201070044)
VEDANT PAWAR (202201040094)

# Content

# Introduction

Grammar correction is a key challenge in Natural Language Processing. Traditional methods often fail to understand complex sentence contexts. This project compares three deep learning models—LSTM/GRU, Bahdanau Attention, and Transformer—for automated grammar correction. We use a dataset of incorrect and corrected sentences for training and evaluation. Our goal is to identify the most accurate and efficient model. Results show that the Transformer outperforms the others in both speed and precision.

# Research Paper Summary

## Goal

Develop a high-performance grammar correction system powered by deep learning.

## Objectives

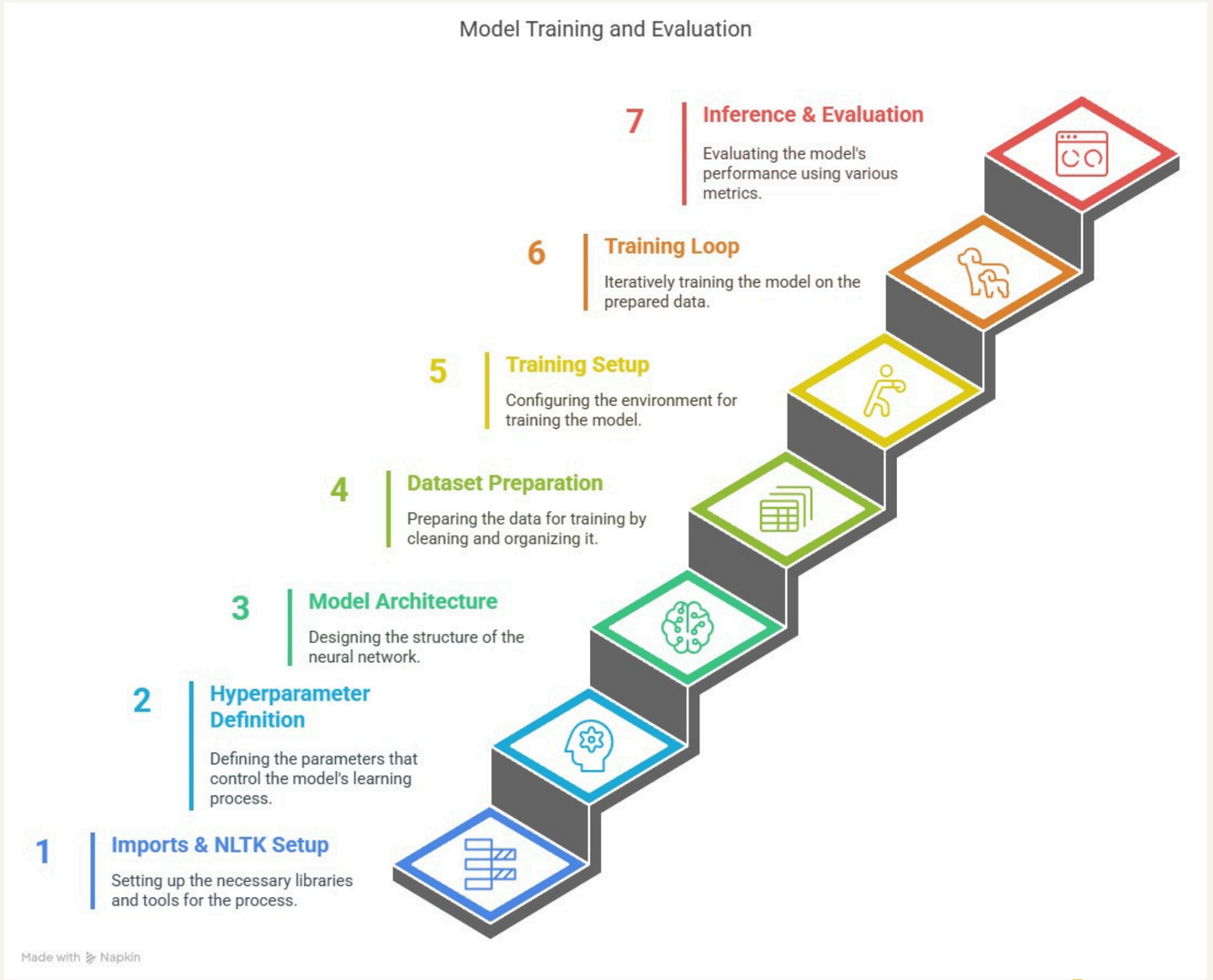
- Evaluate the limitations of traditional grammar tools
- Build an encoder-decoder-based model
- Integrate attention to enhance context understanding
- Validate with real-world datasets

## Problem

Rule-based grammar systems lack adaptability.
We propose data-driven DL models that understand sentence context and structure.

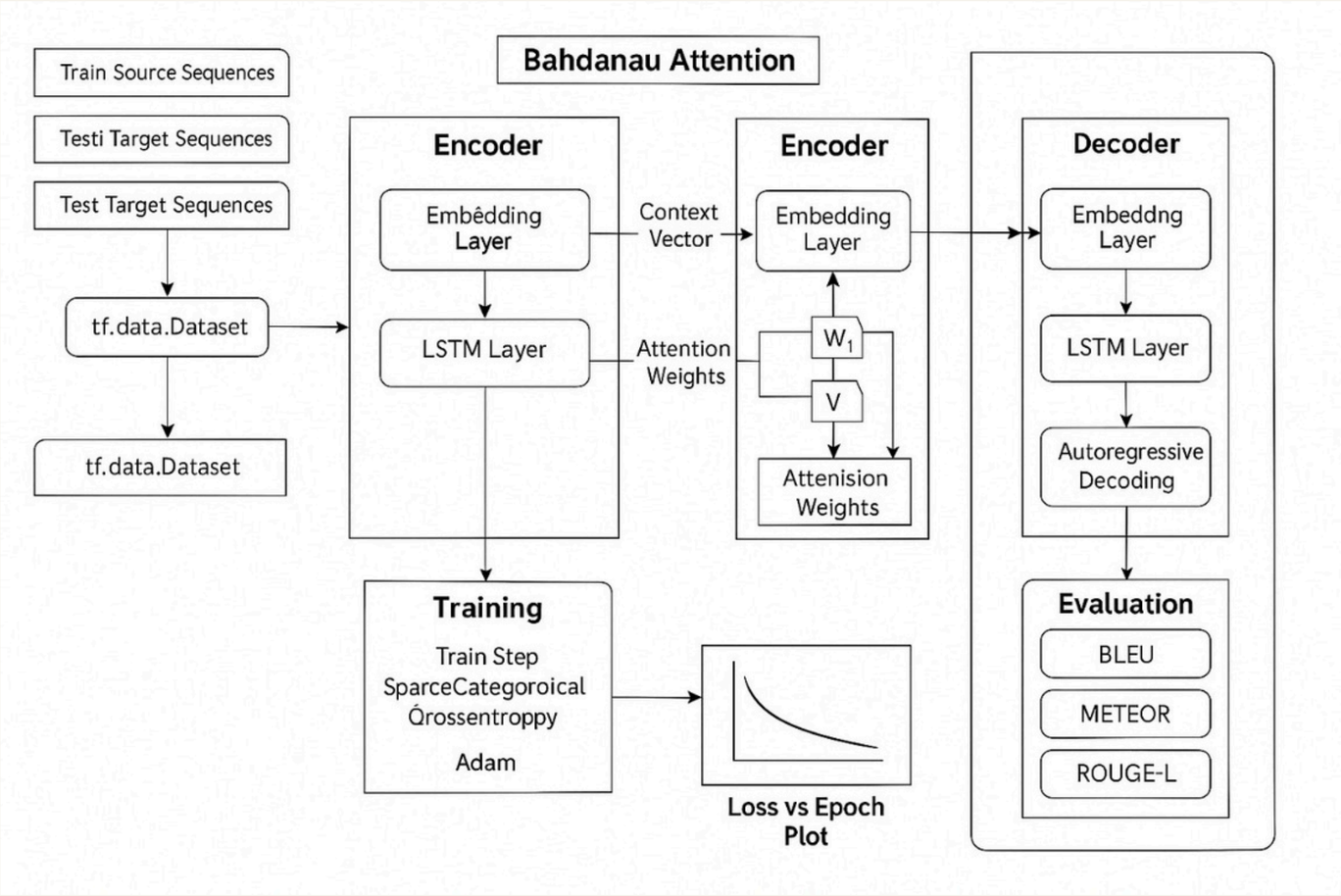
# Model diagram

## Without Attention:



# Model diagram

## With Attention:



# Model Architecture Summary

## Encoder-Decoder Setup:

- Transformer architecture with context-rich encoder (BERT)
- Decoder initialized from scratch

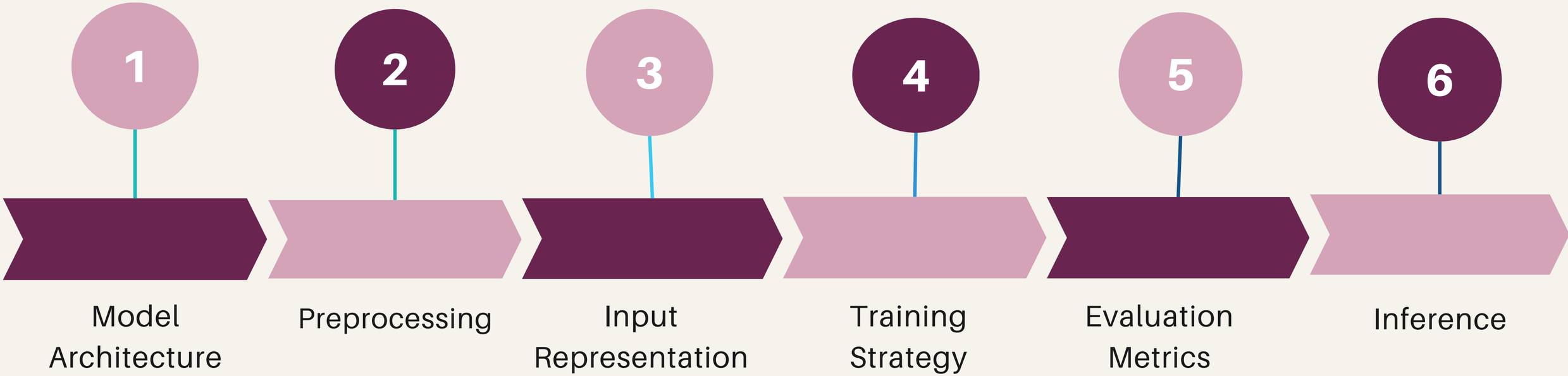
## Input:

- Contextualized tokens from BERT
- Shifted output tokens for decoding

## Preprocessing:

- WordPiece tokenization
- Formatting for BERT ([CLS], [SEP])

# Methodology



# Methodology

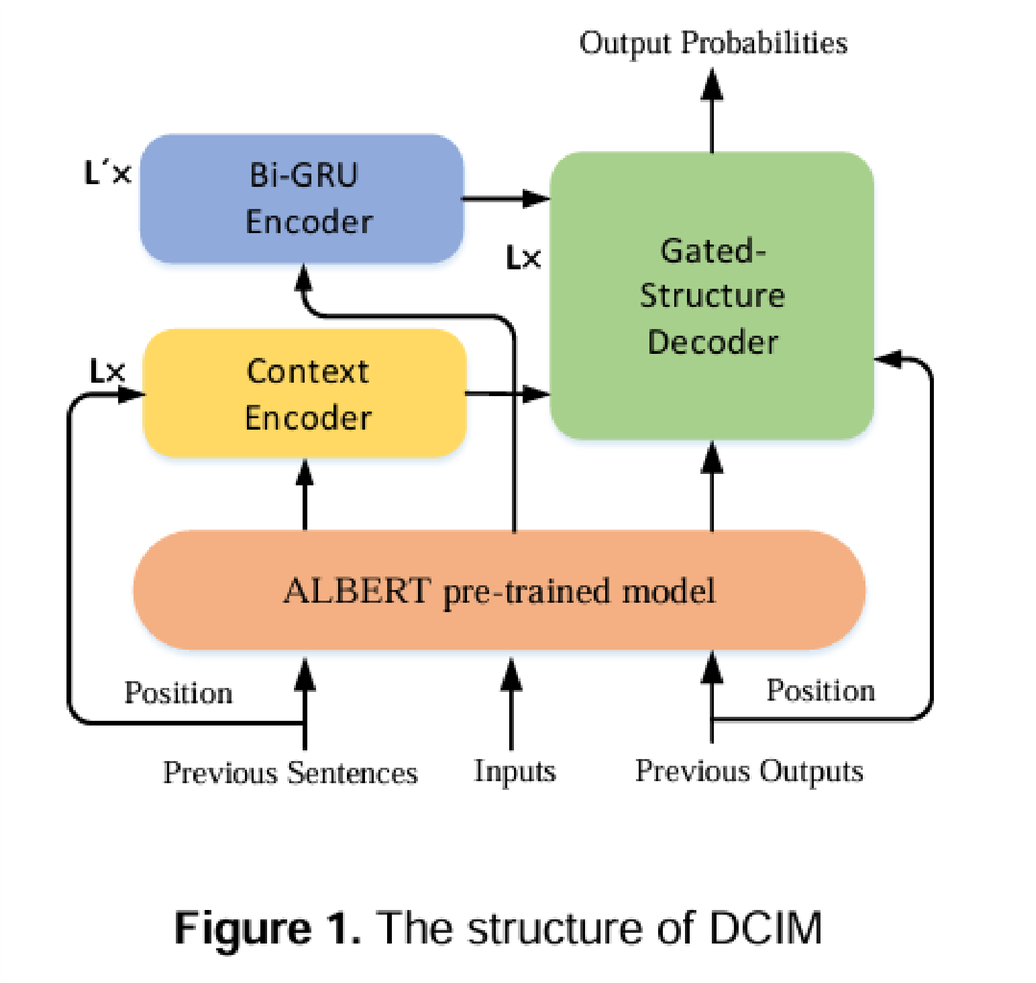



**Figure 1.** The structure of DCIM

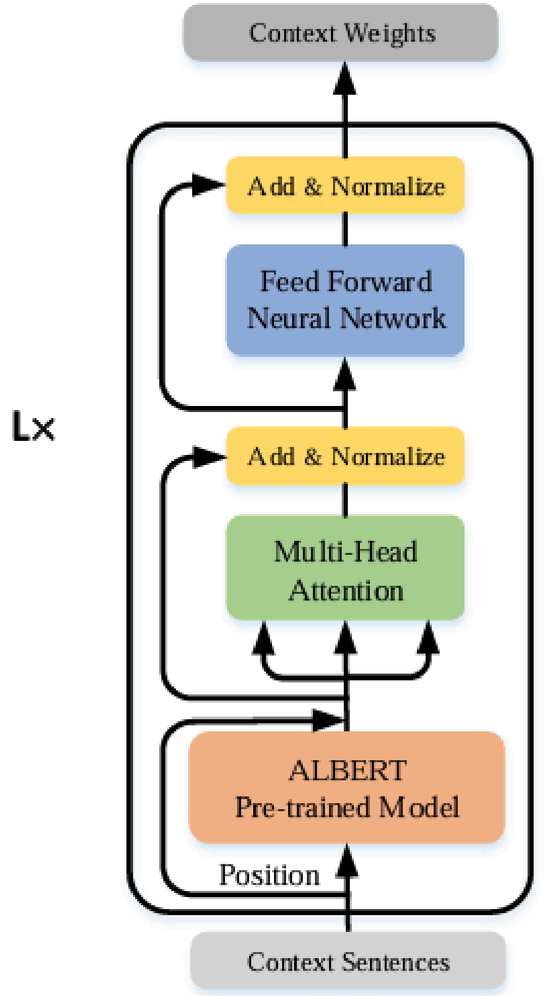



**Figure 2.** The structure of context information encoder

# Dataset Description

**Dataset Name:** Grammar Correction Dataset

**Format:** CSV (Comma-Separated Values)

**Purpose:** Created for training and evaluating grammar correction models in Natural Language Processing (NLP).

**Total Records:** 10,000

**Columns:**

Error Type - Category of the grammatical mistake.

Ungrammatical Statement - The incorrect sentence containing grammar issues.

Standard English - The corrected version of the sentence.