

Figure 1. from
Yuan-Sen Ting et al 2026 Res. Notes AAS 10 doi:10.3847/2515-5172/ae8cf8
https://dx.doi.org/10.3847/2515-5172/ae8cf8
© 2026. The Author(s). Published by the American Astronomical Society.

NASA COSMIC ORIGINS · AI/ML STIG

AI & Machine Learning for Astrophysics
Textbook Edition

PART I FOUNDATIONS

01 The Future of AI and the Mathematical and Physical Sciences
02 PyTorch and Autodifferentiation
03 JAX
04 Inductive Biases

PART II DEEP LEARNING ARCHITECTURES

05 Convolutional Neural Networks
06 Recurrent Neural Networks
07 **Transformers**
08 Graph Neural Networks
09 Equivariant Neural Networks I
10 Equivariant Neural Networks II

PART III GENERATIVE MODELS & INFERENCE

11 Normalizing Flows
12 Diffusion Models
13 Flow Matching
14 Simulation-Based Inference

PART II · DEEP LEARNING ARCHITECTURES · CHAPTER 7

Transformers

Based on the lecture by **Helen Qu** · Flatiron Institute · Interactive notebook

LEARNING OBJECTIVES

- Understand what self-attention computes as a sequence-to-sequence mixing operator.
- Formulate the mathematics of scaled dot-product attention, causal masking, and positional encoding.
- Implement a GPT-style Transformer block from scratch in PyTorch.
- Train a small autoregressive language model on character-level data.
- Sample from the trained model to generate novel text.

Introduction

In Chapter 6, we explored Recurrent Neural Networks (RNNs) and saw how they process sequential data by maintaining a hidden state that updates step-by-step. We showed how a model can generate

ON THIS PAGE

Introduction

- Setting Up the Environment

Preparing the Data

- Loading the Dataset
- Building the Character Vocabulary
- Creating the Train and Validation Splits
- Constructing the Language Modeling Dataset

Building the Model

- Understanding Transformer Building Blocks
- Computing Masked Self-Attention
- Assembling the GPT-Style Model

Training the Model

- Tracking Training Metrics

Evaluating the Results

- Generating Text via Sampling

Summary

Practice Exercises

Further Reading