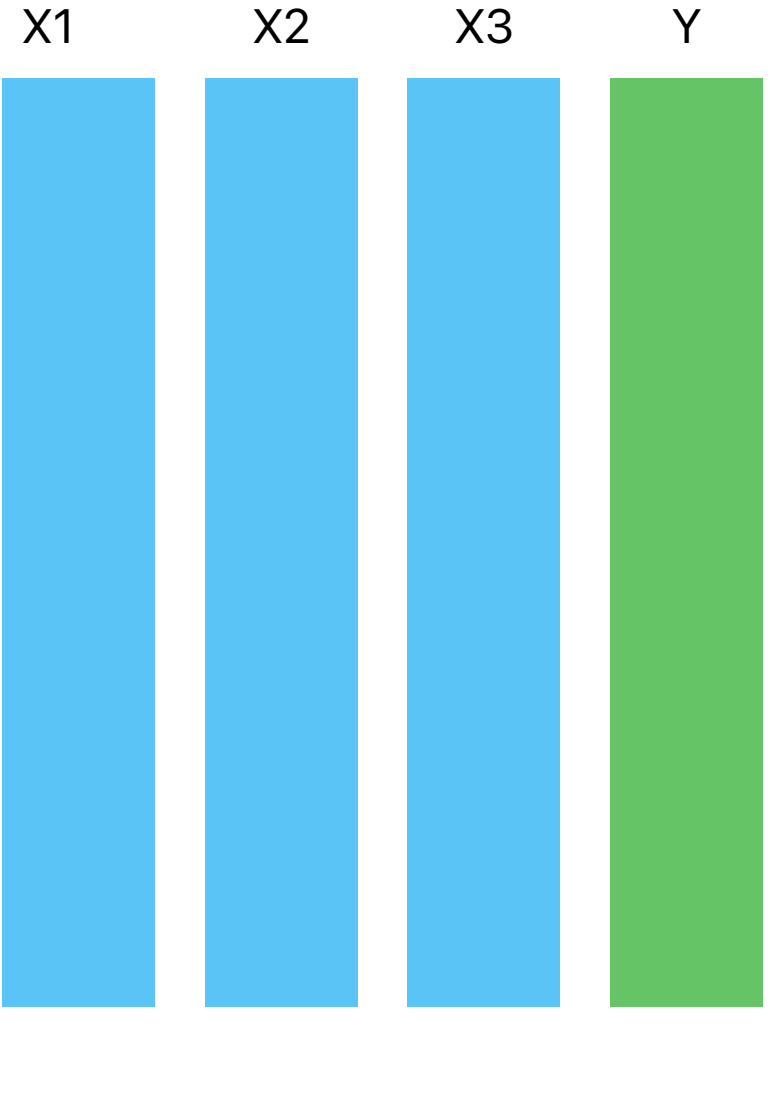


Features in a numerical format



Learn the patterns between input and output

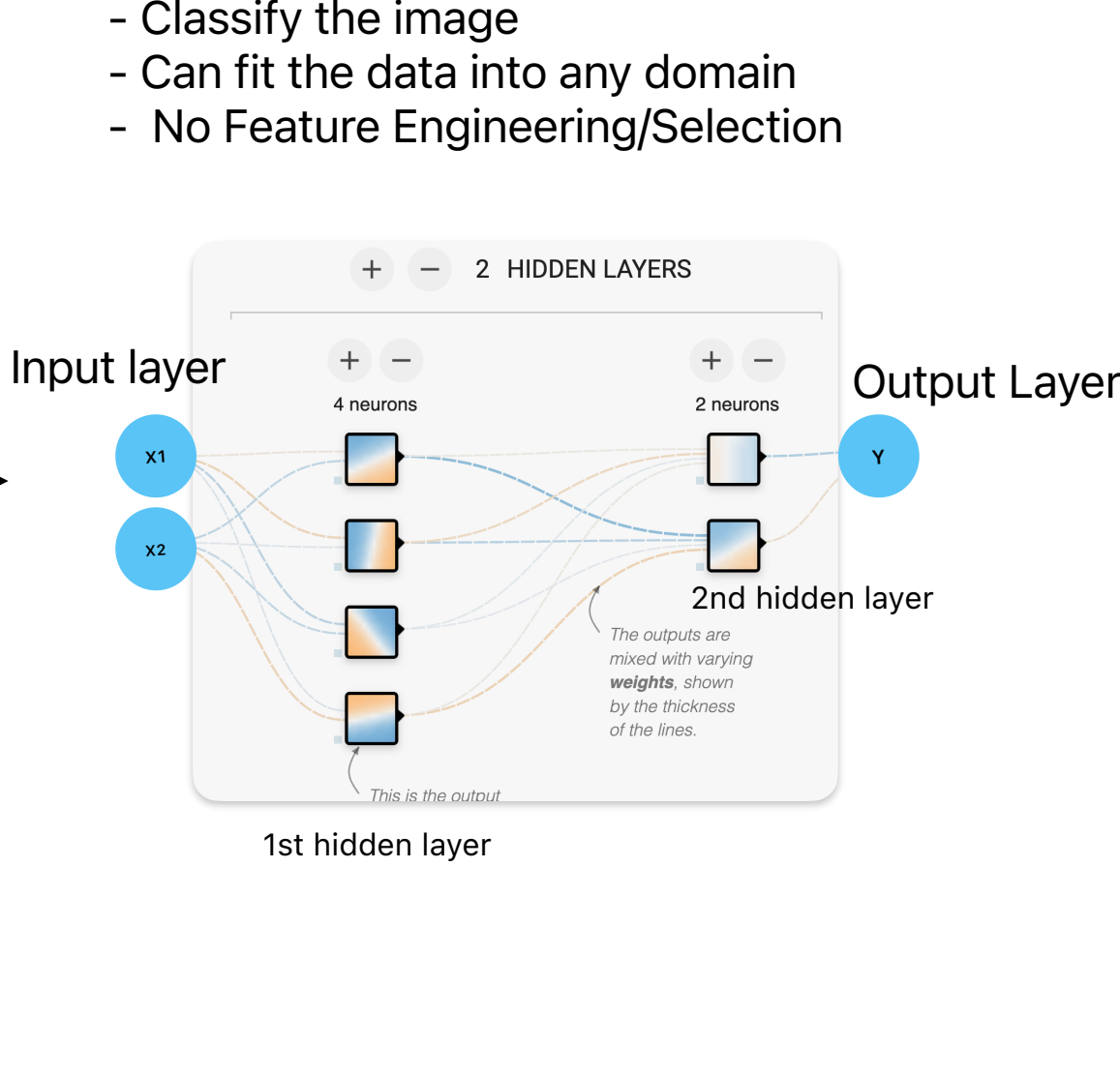


Logistic Regression
Decision Tree Classifier
Naive Bayes
KNN Classifiers
Random Forest - boot strapping + egg
Bagging
Bosting

Data is linearly separable

Deep Learning:
Subset of ML - Artificial Neural Networks → X & Y

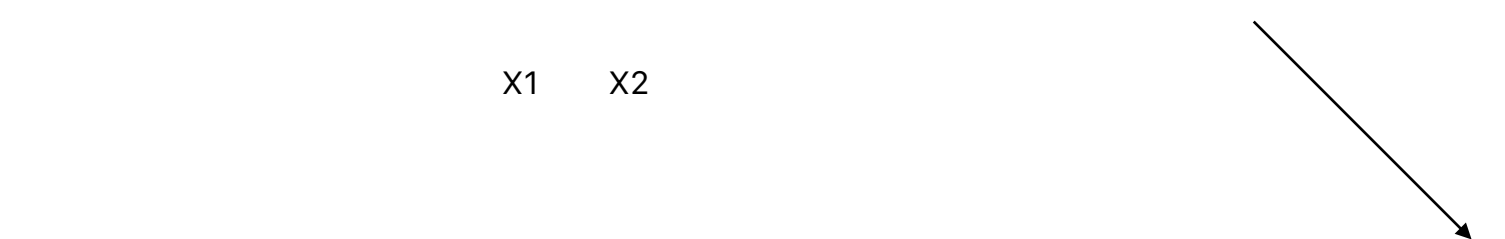
Features in a numerical format



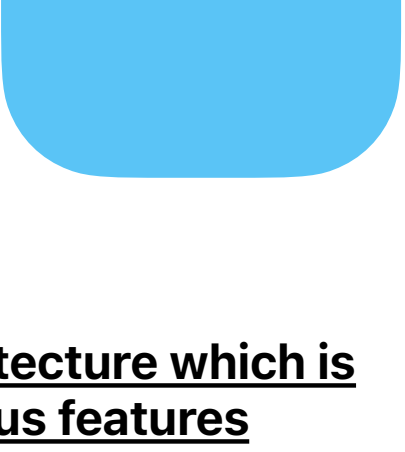
Advantages:

- Can now learn the complex pattern in dataset
- Classify the image
- Can fit the data into any domain
- No Feature Engineering/Selection

Challenges of NLP with Deep Learning Neural Network(Feed Forward Neural Network)



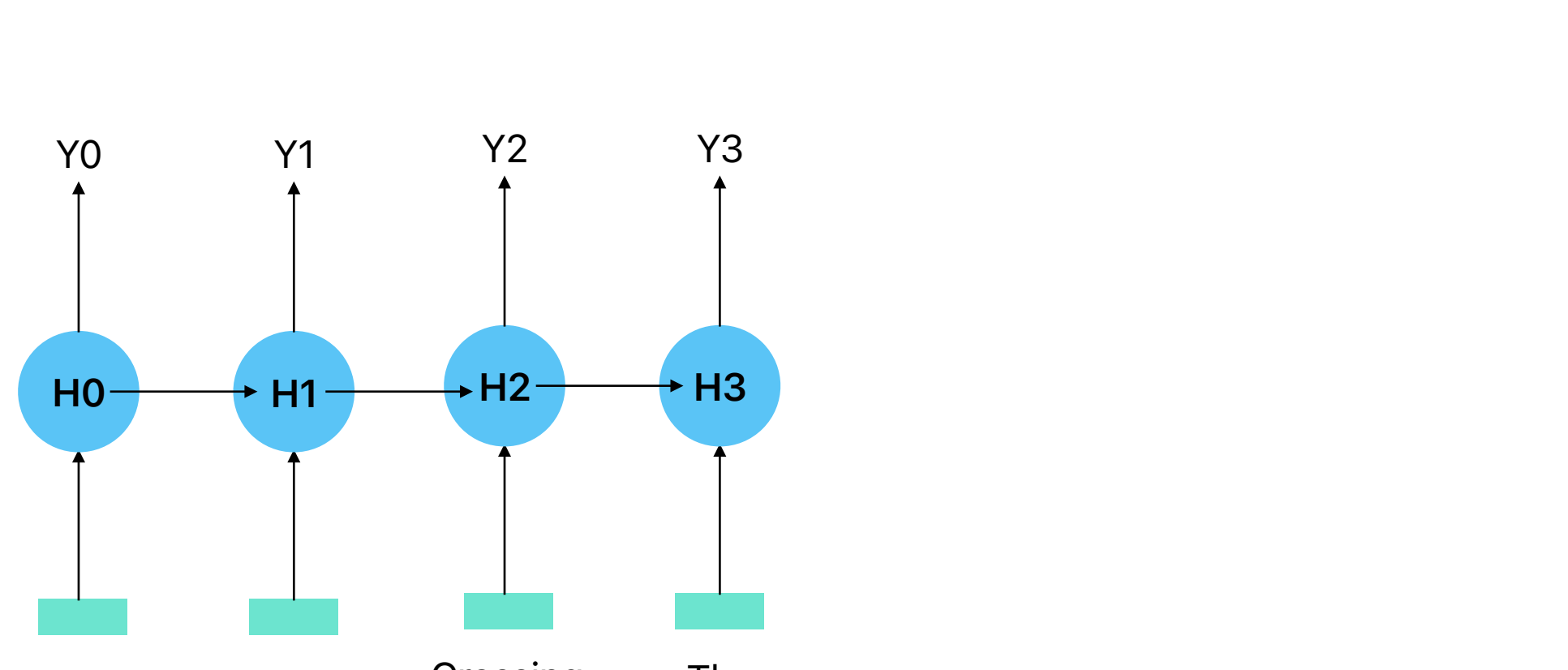
Each feature will be independent



Goal : Find the Neural Network architecture which is capable of remembering previous features

Recurrent Neural Network

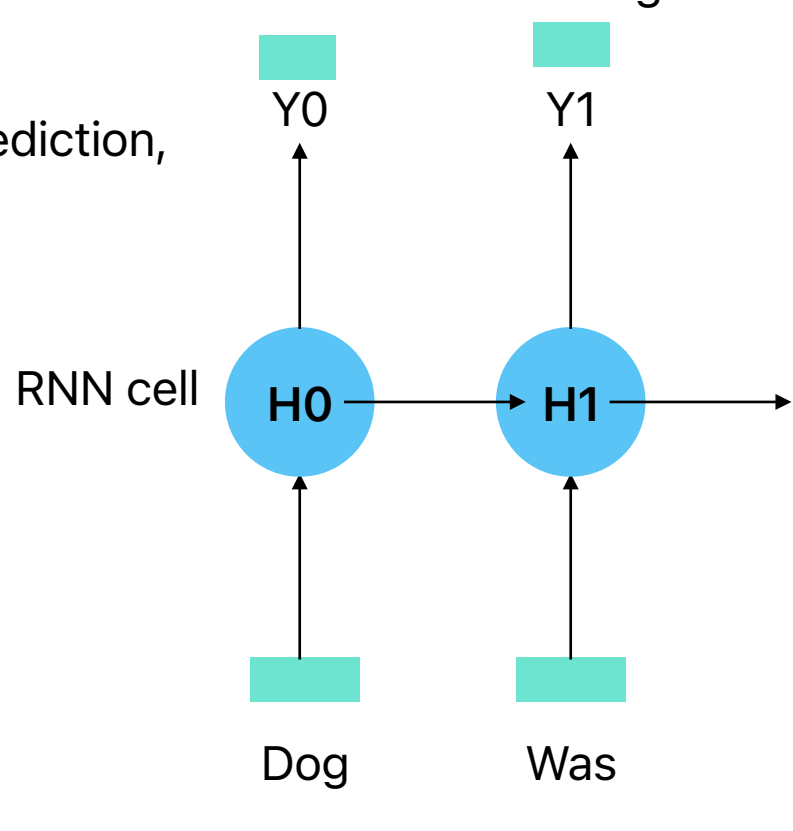
in case of FFNN ----> Features are processed independently



Types of Recurrent Neural Network

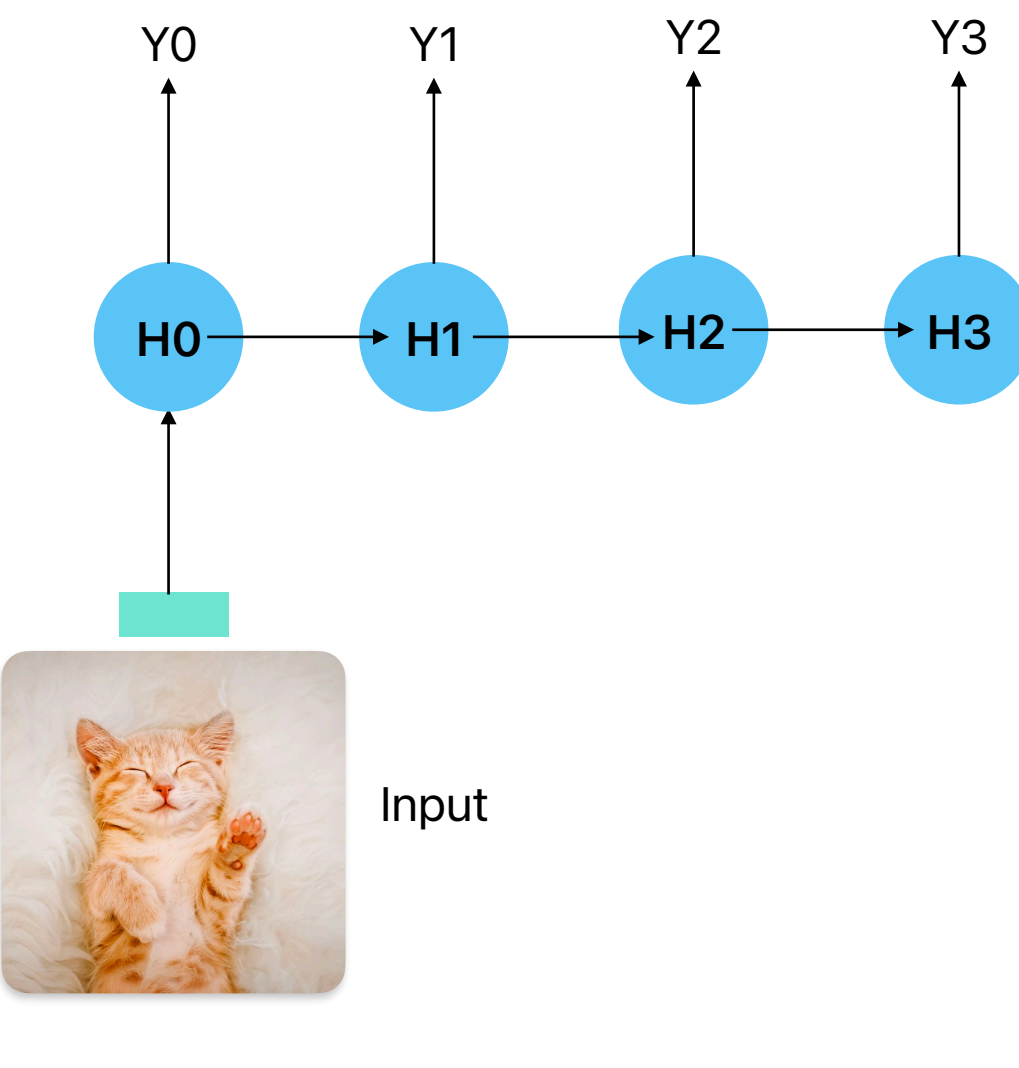
1. One to One Architecture

Application: Next Word Prediction, Next Character Prediction,



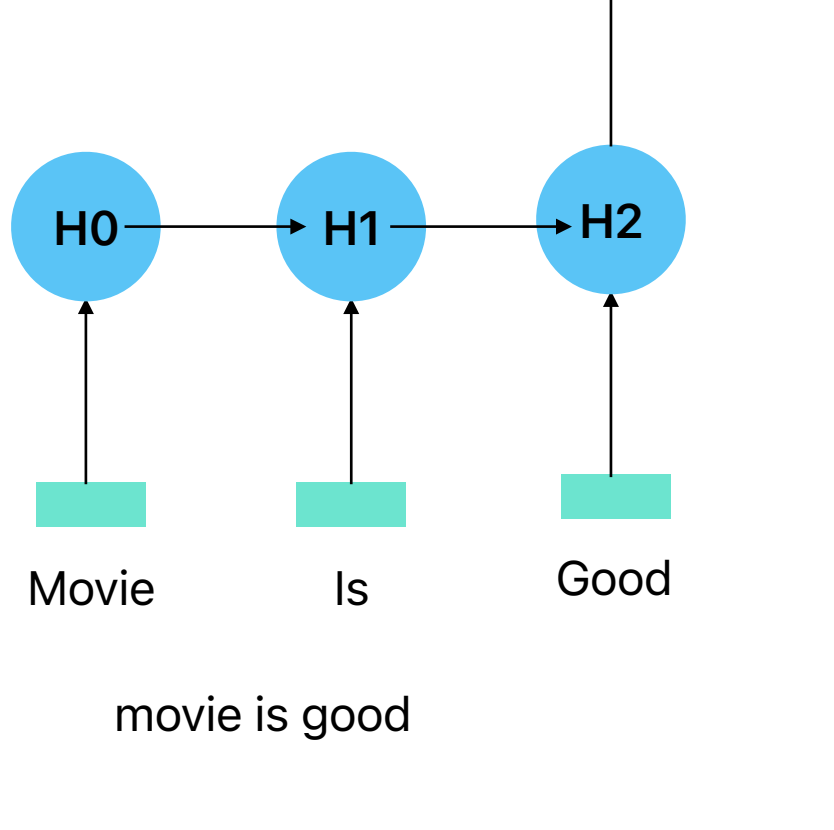
2. One to Many Architecture

Application: Image Captioning



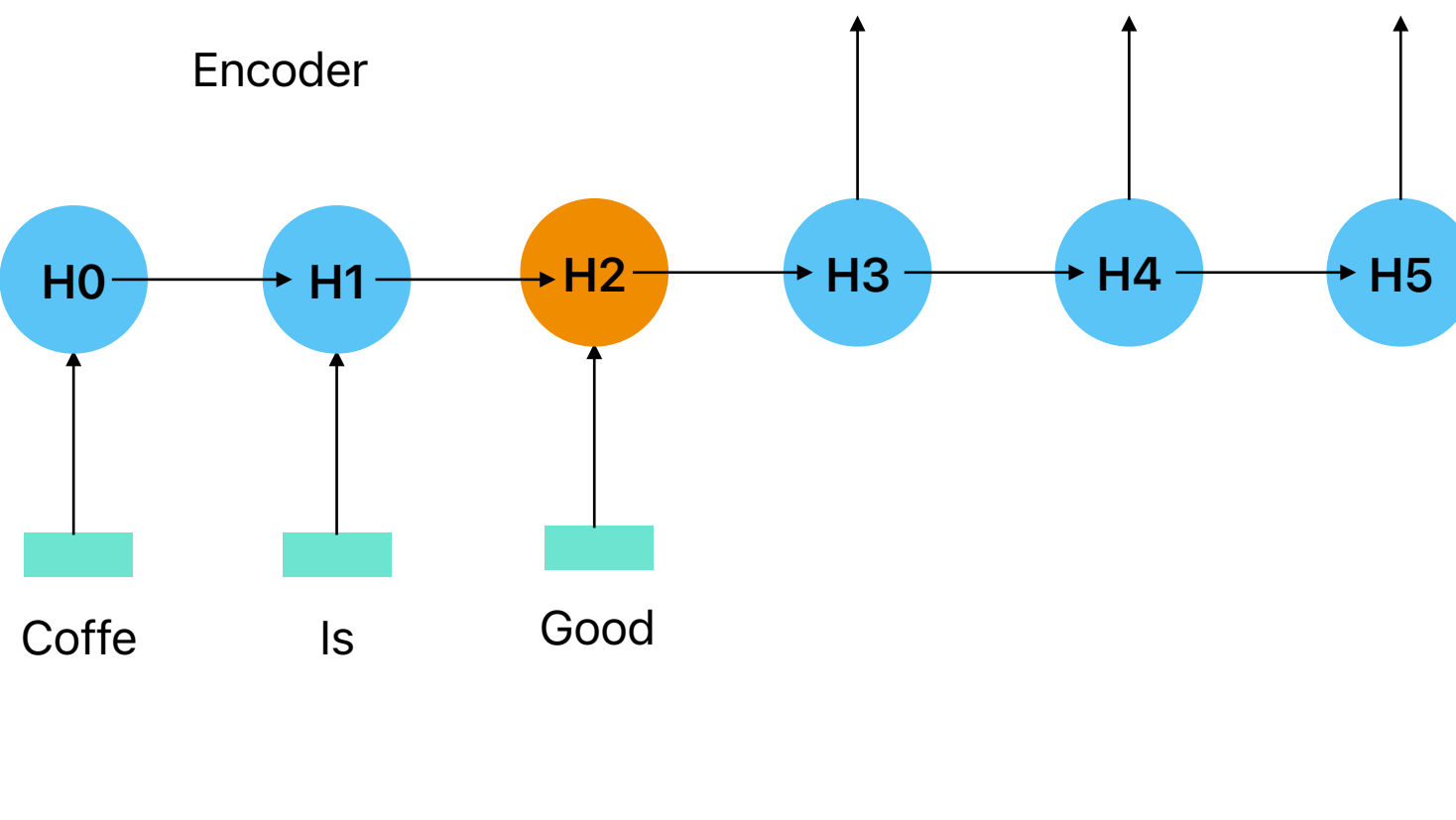
3. Many to One Architecture

Application : Sentiment Classification



4. Many to Many Architecture

Application : Machine Translation

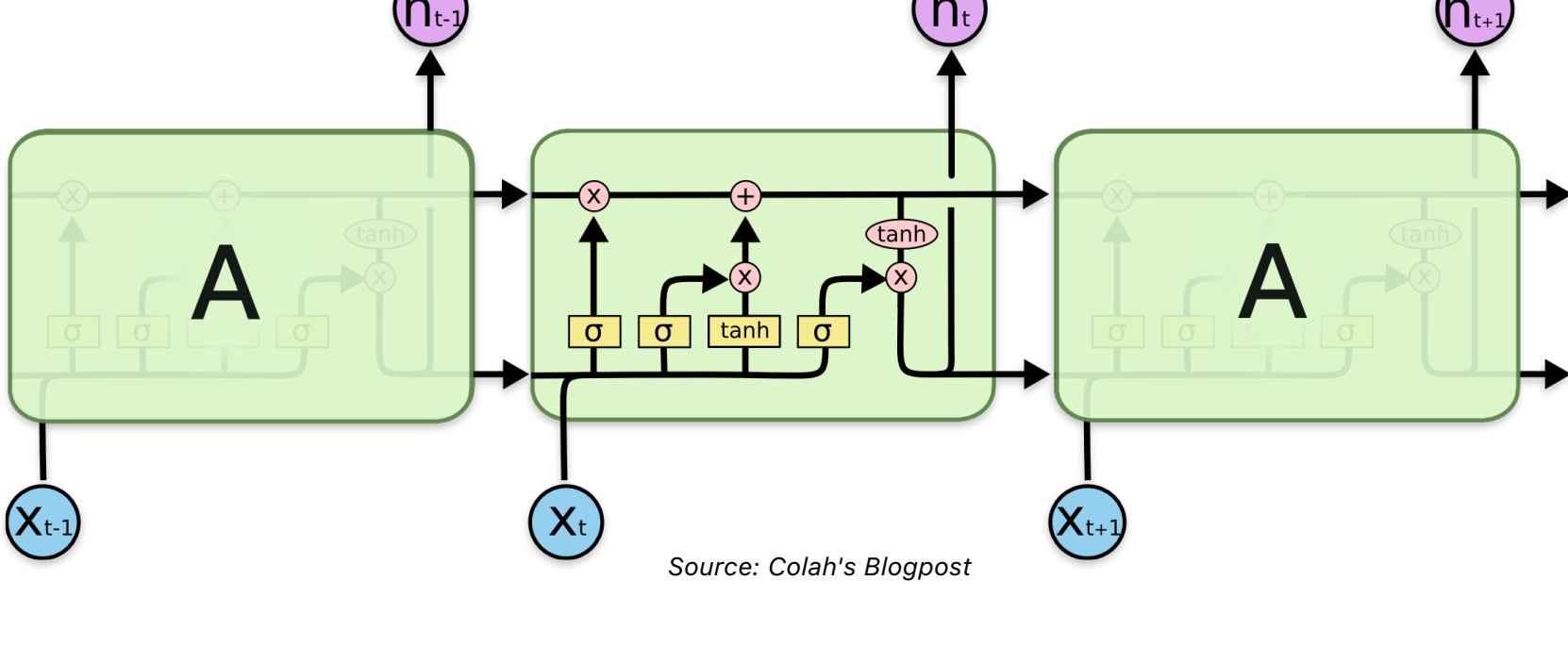


Issues in RNN :

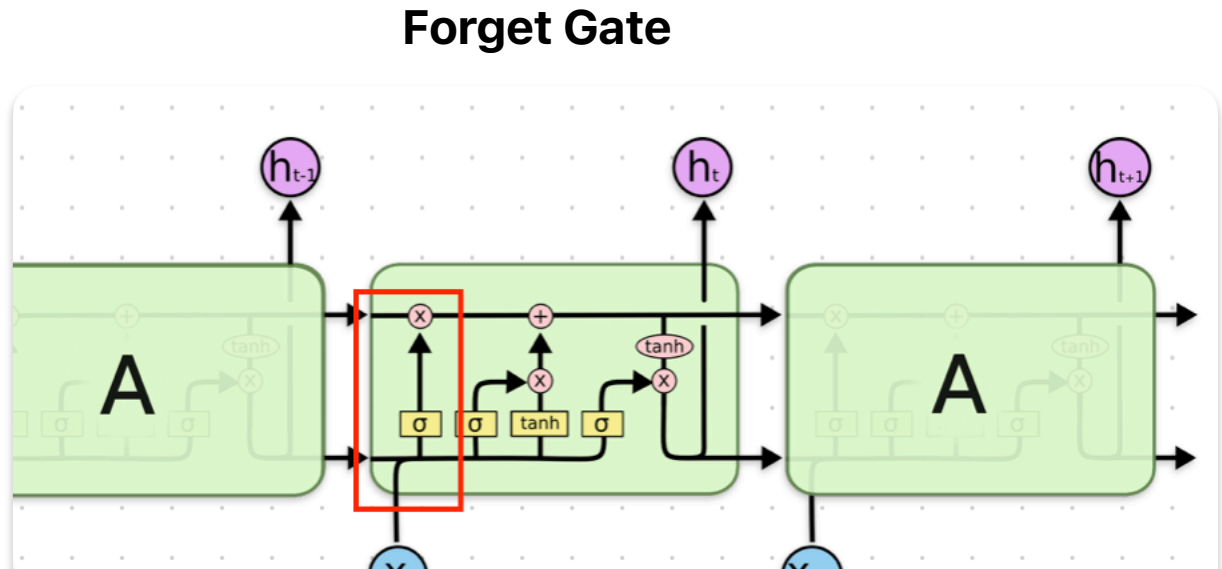
1. Vanishing Gradient Problem
2. Exploding Gradient Problem

- Initialisation , Usage of Proper Activation Function, Gradient Clipping , LSTM
- Initialisation , Truncated BPP, Gradient Clipping

LSTM Long Short Term Memory networks



Forget Gate



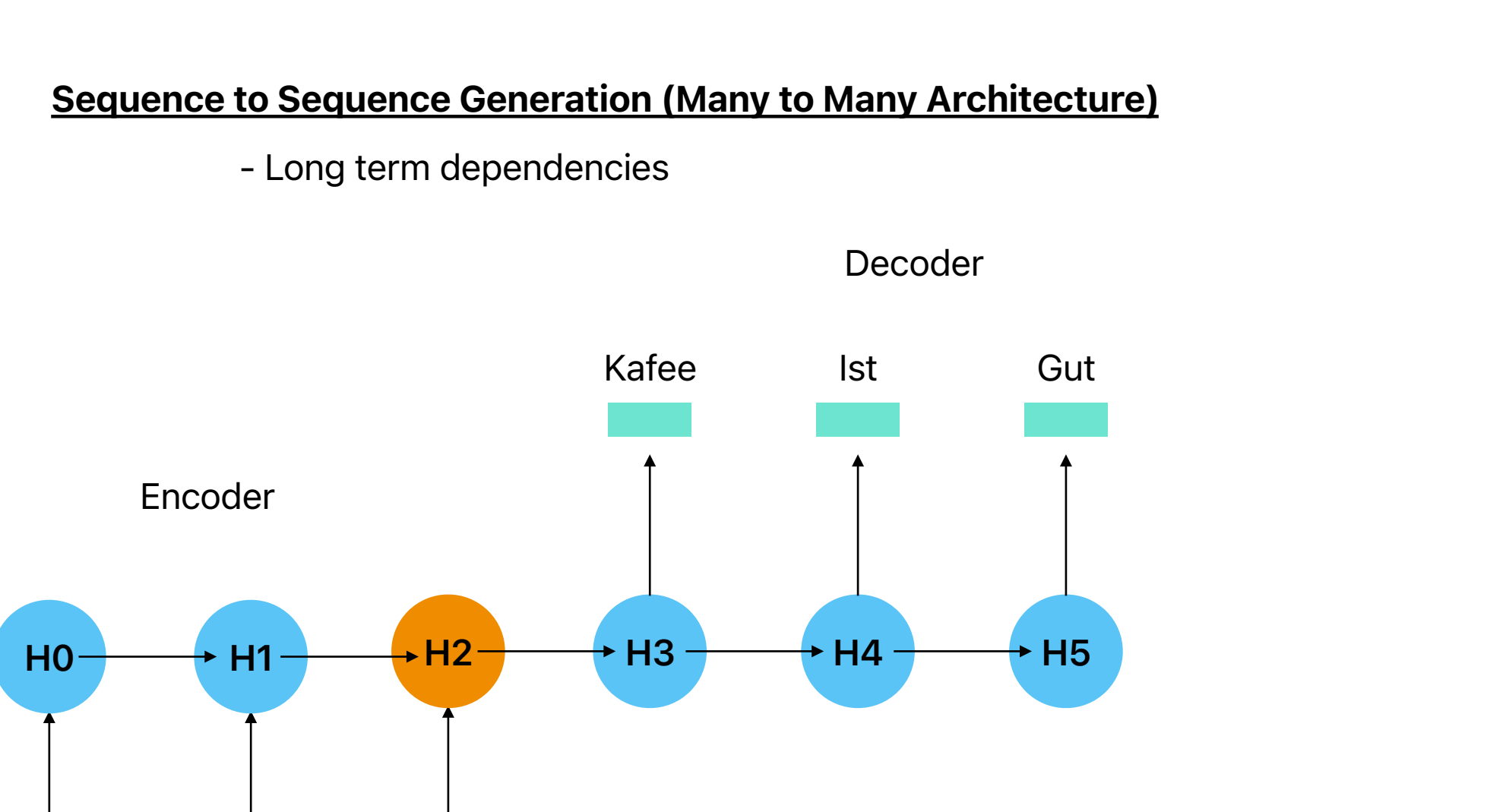
Input Gate

Output Gate

Dog was crossing the Road, and it was very tired

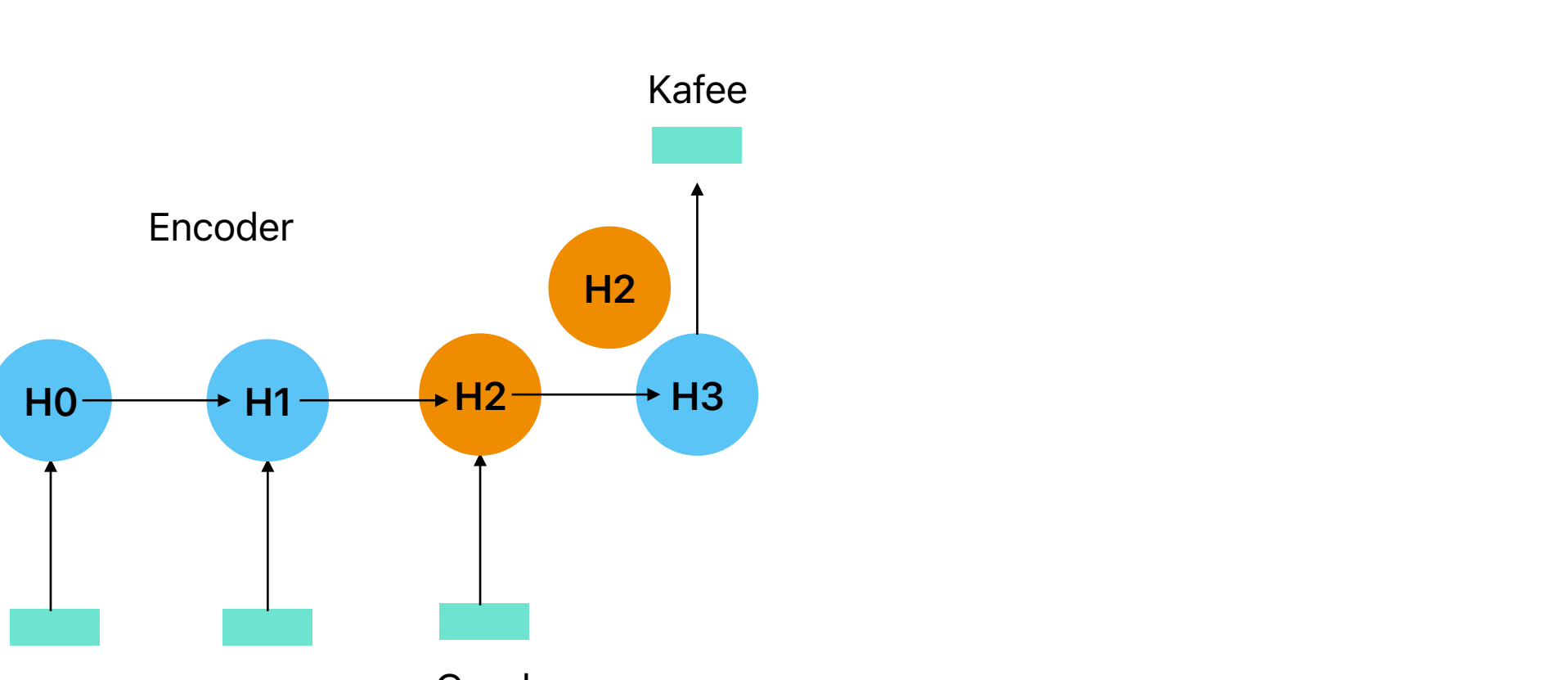
Sequence to Sequence Generation (Many to Many Architecture)

- Long term dependencies

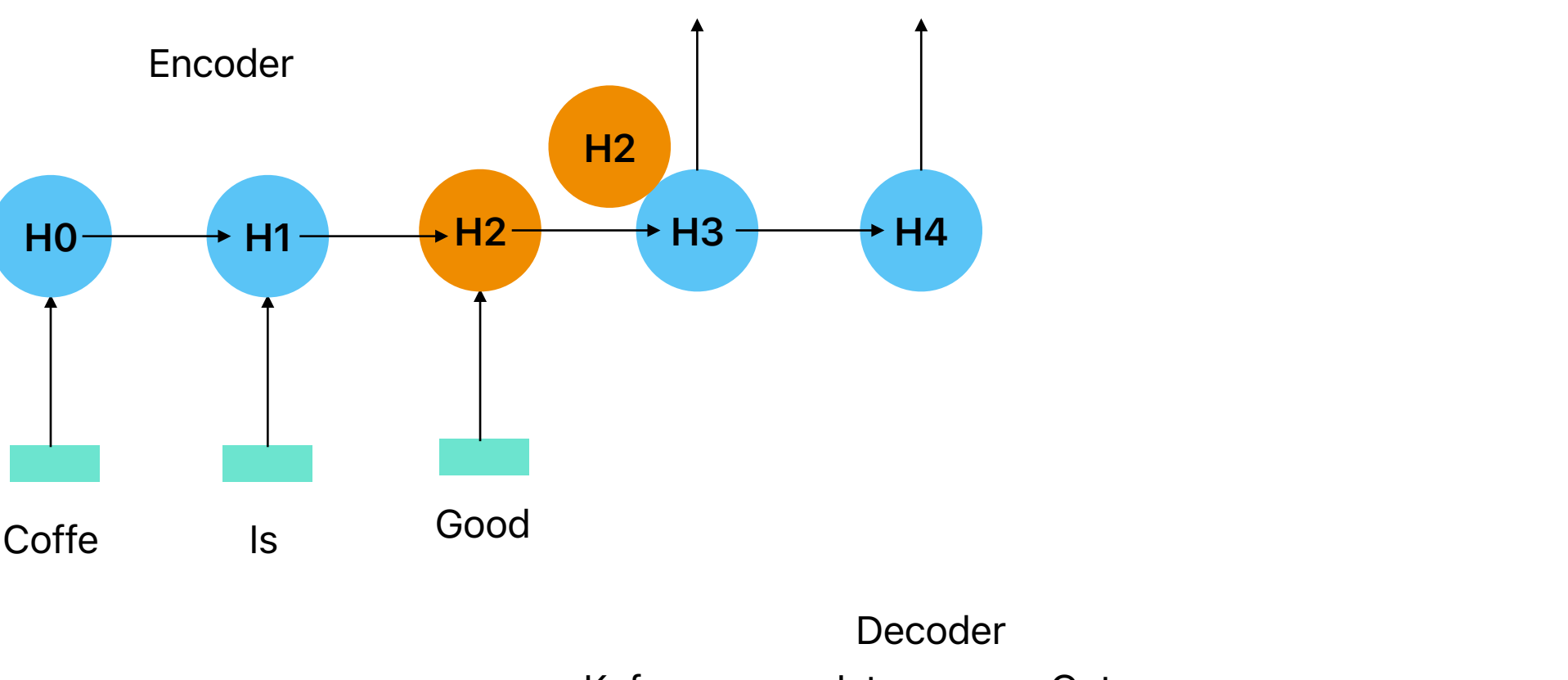


Attention Mechanism (Attention is All you Need)

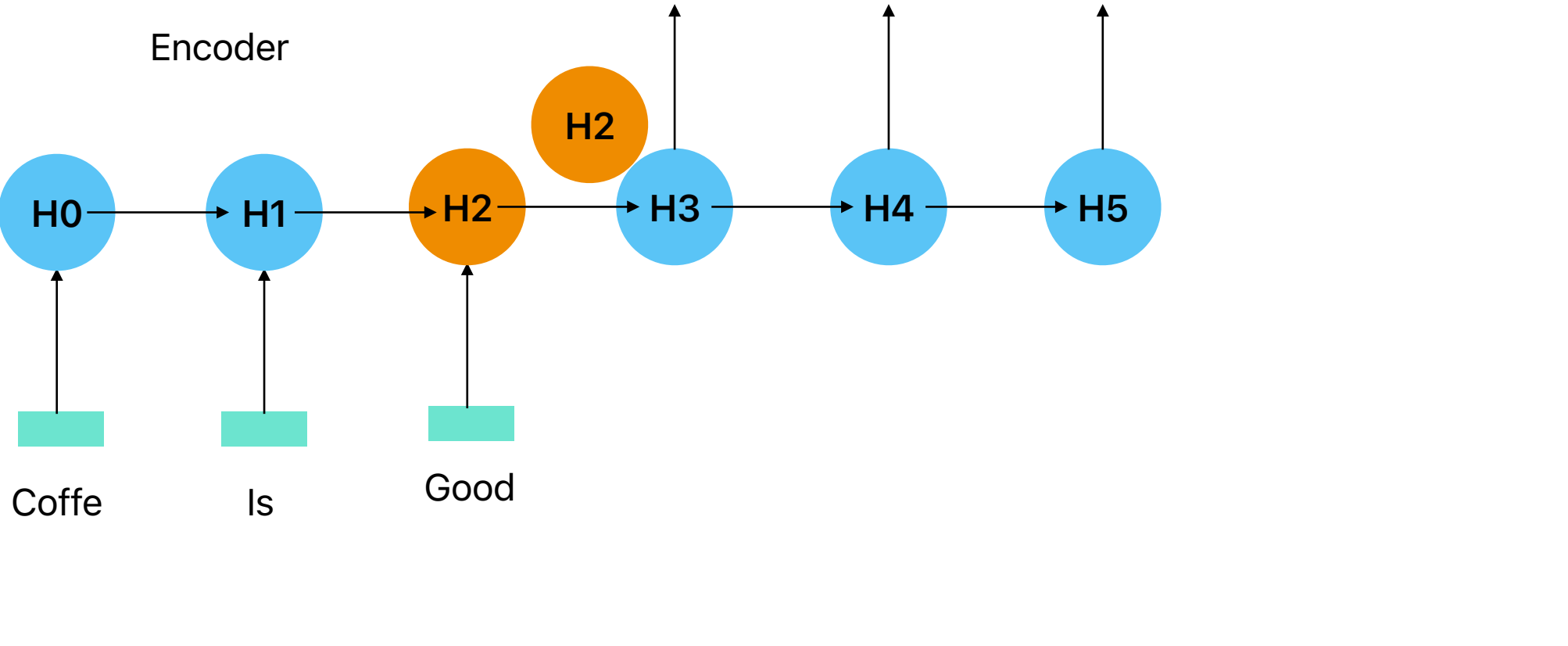
Step 1 of Decoder



Step 2 of Decoder



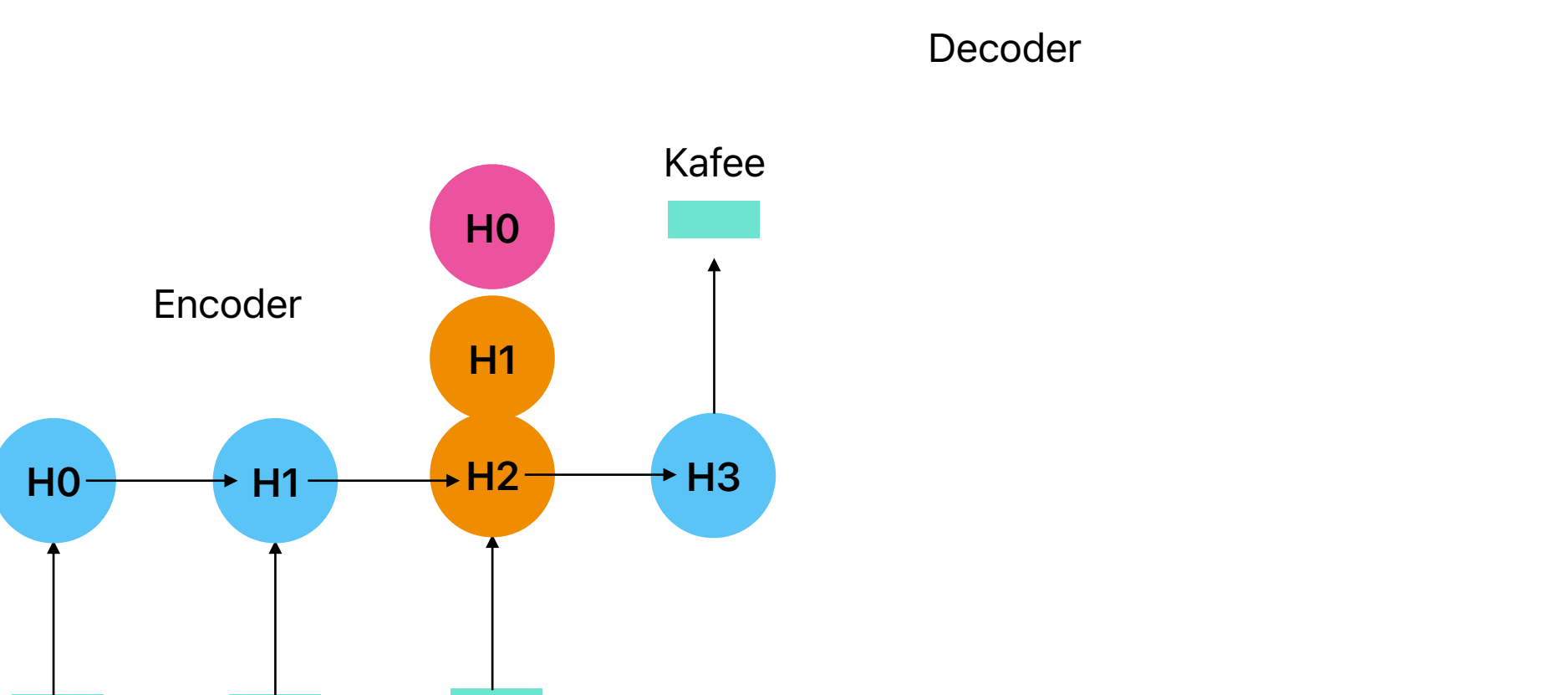
Step 3 of Decoder



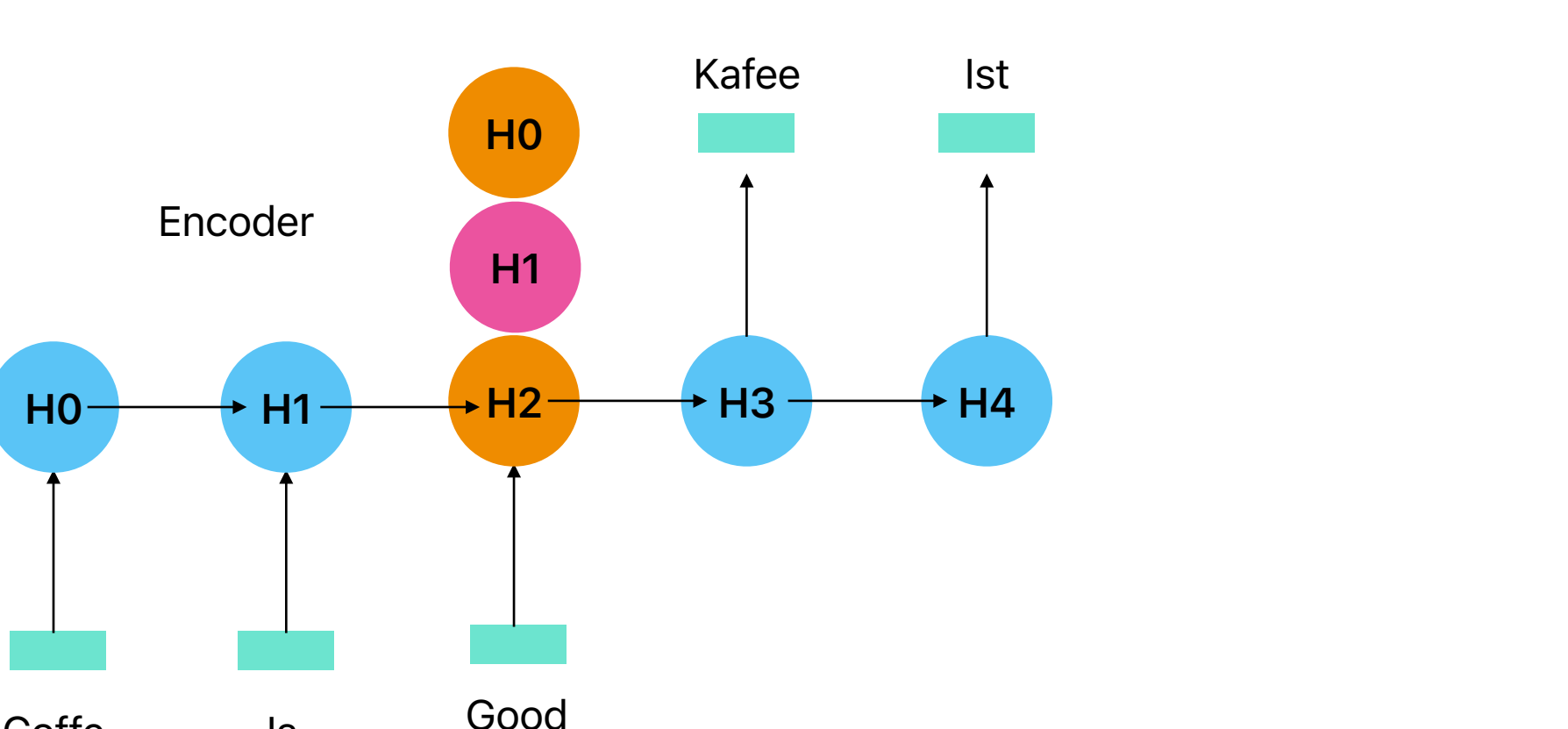
With Attention Mechanism

Decoder

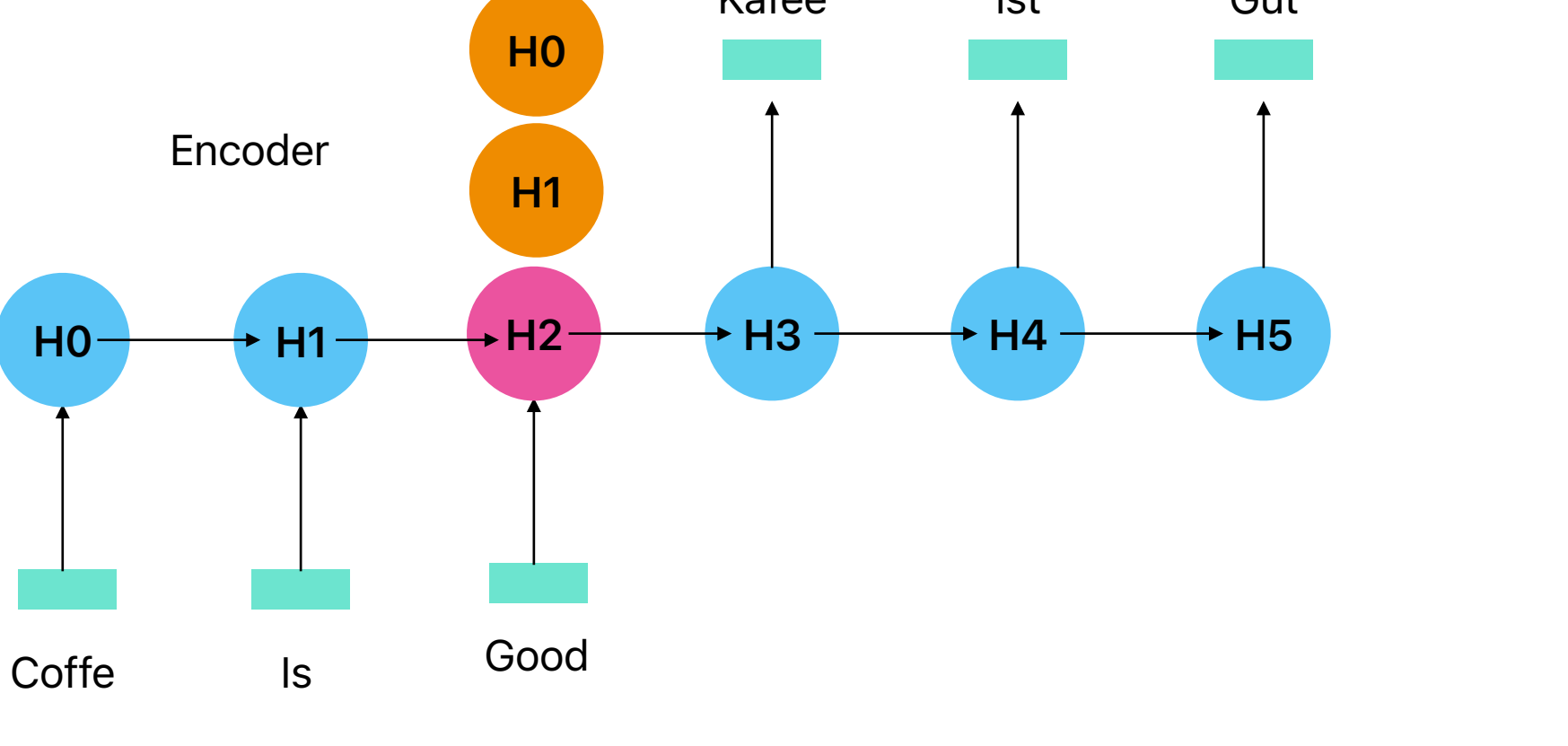
Step 1 of Decoder



Step 2 of Decoder

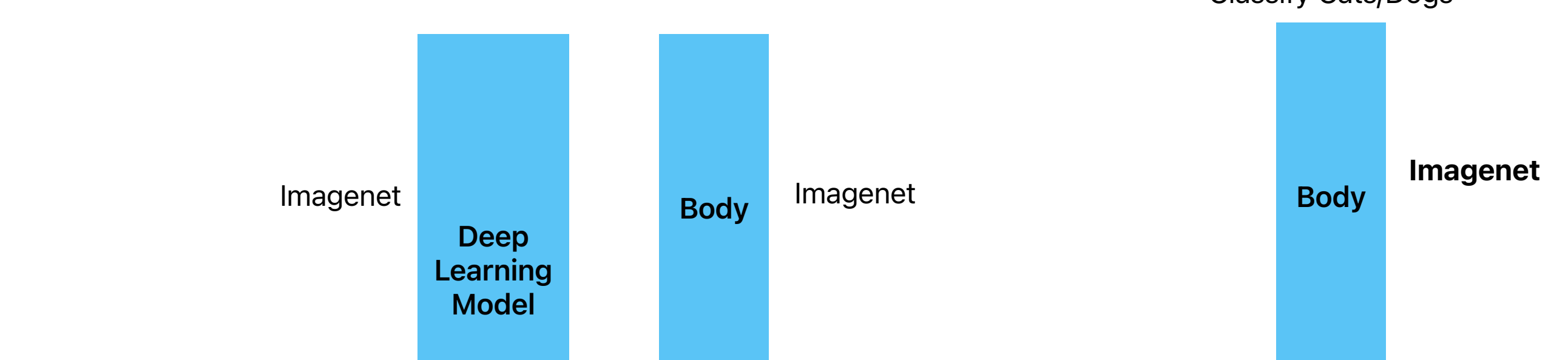


Step 3 of Decoder



Transfer Learning

Computer Vision --> Pretrained Neural Network in ResNet
Adapt with our custom data



1. Pretraining :

Consider a large corpus.
Build/train a model on large corpus with Dummy task of given a current word, what will be the next word.
Language Modelling

2. Domain Adaptation:

Fit on domain data

3. Fine tuning:

- Add a classification layer
- take care of task at hand