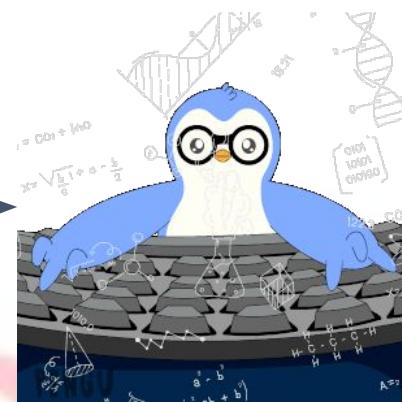


TECHCRUSH ARTIFICIAL INTELLIGENCE BOOTCAMP

Facilitator: Hammed Obasekore
August 15th, 2025

Recap



Programming



Disclaimer: This training material belongs to techcrush and shouldn't be shared

Is Running Others Code Relevant to My AI Journey?



Generative
pre-trained
transformer

Disclaimer: This training material belongs to techcrush and shouldn't be

Attention Is All You Need

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
jakob@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Illia Polosukhin*
illia.polosukhin@google.com

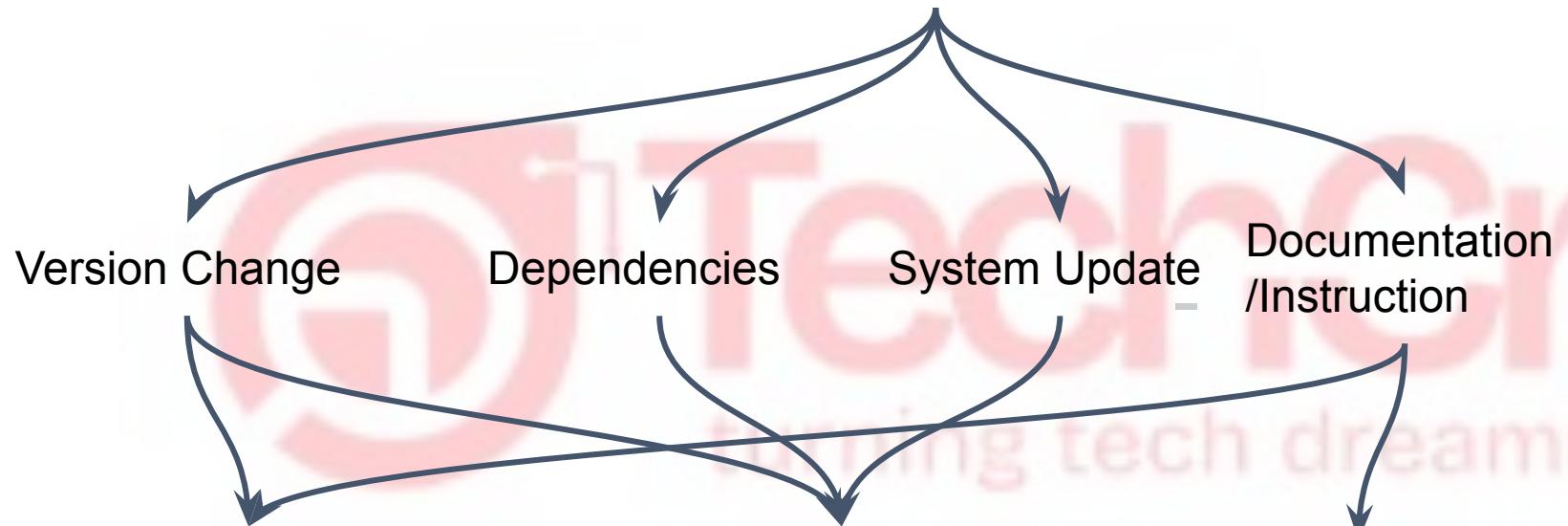
The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 41.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new state-of-the-art BLEU score of 41.8 after training in 3.5 days, eight times a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with coarse and limited training data.

*Equal contribution. Listing order is random. Jakob proposed replacing RNNs with self-attention and started the effort to implement this idea. Ashish and Illia designed and implemented the first Transformer models and has been crucially involved in every aspect of this work. Noam proposed scaled dot-product attention, multi-head attention and the position-wise feed-forward representation and became the other person involved in nearly every detail. Niki designed, implemented, tuned and evaluated countless model variants in our original codebase and tensor2tensor. Llion also experimented with novel model variants, was responsible for our initial codebase, and efficient inference and visualizations. Lukasz and Aidan spent countless long days designing various parts of and implementing tensor2tensor, replacing our earlier codebase, greatly improving results and massively accelerating our research.

†Work performed while at Google Brain.

‡Work performed while at Google Research.

Debugging is An Art



Lemon-Quality
-Classification

Pose Landmarker style transfer

90% OF
PROGRAMMING
IS CODING.
THE OTHER
90% IS
DEBUGGING

Python

- Interpreted language
- Syntax & Semantic
 - White-Space and Indentation
- Comments
- Variables & Data Types
 - Variable Name
 - Case-Sensitivity
 - Single or Double quotes
 - Data-mismatch
 - Scope (Global & Local)
- Array-like
 - Strings
 - Lists
 - Tuples
 - sets
 - Dictionaries
- Operator
- Condition
- Loops
- Functions
- Importing and Installing libraries

[Visualize Python Execution](#)

[W3School - Python](#)



Disclaimer: This training material belongs to techcrush and shouldn't be shared