Title: Feedback neural network

URL: https://en.wikipedia.org/wiki/Feedback_neural_network

PageID: 79132280

Categories: Category:Artificial intelligence, Category:Large language models



---

Feedback neural network are neural networks with the ability to provide bottom-up and top-down design feedback to their input or previous layers, based on their outputs or subsequent layers. This is notably used in large language models specifically in reasoning language models (RLM). This process is designed to mimic self-assessment and internal deliberation, aiming to minimize errors (like hallucinations ) and increase interpretability . Reflection is a form of " test-time compute ", where additional computational resources are used during inference.

## Introduction

Traditional neural networks process inputs in a feedforward manner, generating outputs in a single pass. However, their limitations in handling complex tasks, and especially compositional ones, have led to the development of methods that simulate internal deliberation. Techniques such as chain-of-thought prompting encourage models to generate intermediate reasoning steps, thereby improving their performance in such tasks.

The feedback can take place either after a full network pass and decoding to tokens, or continuously in latent space (the last layer can be fed back to the first layer). [ 1 ] [ 2 ] In LLMs, special tokens can mark the beginning and end of reflection before producing a final response (e.g., ).

This internal process of "thinking" about the steps leading to an answer is designed to be analogous to human metacognition or "thinking about thinking". It helps AI systems approach tasks that require multi-step reasoning, planning, and logical thought.

## Techniques

Increasing the length of the Chain-of-Thought reasoning process, by passing the output of the model back to its input and doing multiple network passes, increases inference-time scaling. [ 3 ] Reinforcement learning frameworks have also been used to steer the Chain-of-Thought. One example is Group Relative Policy Optimization (GRPO), used in DeepSeek-R1 , [ 4 ] a variant of policy gradient methods that eliminates the need for a separate "critic" model by normalizing rewards within a group of generated outputs, reducing computational cost. Simple techniques like "budget forcing" (forcing the model to continue generating reasoning steps) have also proven effective in improving performance. [ 5 ]

## Types of reflection

### Post-hoc reflection

Analyzes and critiques an initial output separately, often involving prompting the model to identify errors or suggest improvements after generating a response. The Reflexion framework follows this approach. [ 6 ]

### Iterative reflection

Revises earlier parts of a response dynamically during generation. Self-monitoring mechanisms allow the model to adjust reasoning as it progresses. Methods like Tree-of-Thoughts exemplify this, enabling backtracking and alternative exploration.

### Intrinsic reflection

Integrates self-monitoring directly into the model architecture rather than relying solely on external prompts, enabling models with inherent awareness of their reasoning limitations and uncertainties.

This has been used by Google DeepMind in a technique called Self-Correction via Reinforcement Learning (SCoRe) which rewards the model for improving its responses. [ 7 ]

Process reward models and limitations

Early research explored PRMs to provide feedback on each reasoning step, unlike traditional reinforcement learning which rewards only the final outcome. However, PRMs have faced challenges, including computational cost and reward hacking. DeepSeek-R1's developers found them to be not beneficial. [ 8 ] [ 9 ]

See also

Reflective programming

Reservoir computing

References

v

t

e

History timeline

timeline

Companies

Projects

Parameter Hyperparameter

Hyperparameter

Loss functions

Regression Bias–variance tradeoff Double descent Overfitting

Bias–variance tradeoff

Double descent

Overfitting

Clustering

Gradient descent SGD Quasi-Newton method Conjugate gradient method

SGD

Quasi-Newton method

Conjugate gradient method

Backpropagation

Attention

Convolution

Normalization Batchnorm

Batchnorm

Activation Softmax Sigmoid Rectifier

Softmax

Sigmoid

Rectifier

Gating

Weight initialization

Regularization

Datasets Augmentation

Augmentation

Prompt engineering

Reinforcement learning Q-learning SARSA Imitation Policy gradient

Q-learning

SARSA

Imitation

Policy gradient

Diffusion

Latent diffusion model

Autoregression

Adversary

RAG

Uncanny valley

RLHF

Self-supervised learning

Reflection

Recursive self-improvement

Hallucination

Word embedding

Vibe coding

Machine learning In-context learning

In-context learning

Artificial neural network Deep learning

Deep learning

Language model Large language model NMT

Large language model

NMT

Reasoning language model

Model Context Protocol

Intelligent agent

Artificial human companion

Humanity's Last Exam

Artificial general intelligence (AGI)

AlexNet

WaveNet

Human image synthesis

HWR

OCR

Computer vision

Speech synthesis 15.ai ElevenLabs

15.ai

ElevenLabs

Speech recognition Whisper

Whisper

Facial recognition

AlphaFold

Text-to-image models Aurora DALL-E Firefly Flux Ideogram Imagen Midjourney Recraft Stable Diffusion

Aurora

DALL-E

Firefly

Flux

Ideogram

Imagen

Midjourney

Recraft

Stable Diffusion

Text-to-video models Dream Machine Runway Gen Hailuo AI Kling Sora Veo

Dream Machine

Runway Gen

Hailuo AI

Kling

Sora

Veo

Music generation Riffusion Suno AI Udio

Riffusion

Suno AI

Udio

Word2vec

Seq2seq

GloVe

BERT

T5

Llama

Chinchilla AI

PaLM

GPT 1 2 3 J ChatGPT 4 4o o1 o3 4.5 4.1 o4-mini 5

1

2

3

J

ChatGPT

4

4o

o1

o3

4.5

4.1

o4-mini

5

Claude

Gemini Gemini (language model) Gemma

Gemini (language model)

Gemma

Grok

LaMDA

BLOOM

DBRX

Project Debater

IBM Watson

IBM Watsonx

Granite

PanGu-Σ

DeepSeek

Qwen

AlphaGo

AlphaZero

OpenAI Five

Self-driving car

MuZero

Action selection AutoGPT

AutoGPT

Robot control

Alan Turing

Warren Sturgis McCulloch

Walter Pitts

John von Neumann

Claude Shannon

Shun'ichi Amari

Kunihiko Fukushima

Takeo Kanade

Marvin Minsky

John McCarthy

Nathaniel Rochester

Allen Newell

Cliff Shaw

Herbert A. Simon

Oliver Selfridge

Frank Rosenblatt

Bernard Widrow

Joseph Weizenbaum

Seymour Papert

Seppo Linnainmaa

Paul Werbos

Geoffrey Hinton

John Hopfield

Jürgen Schmidhuber

Yann LeCun

Yoshua Bengio

Lotfi A. Zadeh

Stephen Grossberg

Alex Graves

James Goodnight

Andrew Ng

Fei-Fei Li

Alex Krizhevsky

Ilya Sutskever

Oriol Vinyals

Quoc V. Le

Ian Goodfellow

Demis Hassabis

David Silver

Andrej Karpathy

Ashish Vaswani

Noam Shazeer

Aidan Gomez

John Schulman

Mustafa Suleyman

Jan Leike

Daniel Kokotajlo

François Chollet

Neural Turing machine

Differentiable neural computer

Transformer Vision transformer (ViT)

Vision transformer (ViT)

Recurrent neural network (RNN)

Long short-term memory (LSTM)

Gated recurrent unit (GRU)

Echo state network

Multilayer perceptron (MLP)

Convolutional neural network (CNN)

Residual neural network (RNN)

Highway network

Mamba

Autoencoder

Variational autoencoder (VAE)

Generative adversarial network (GAN)

Graph neural network (GNN)

Category