Understanding Catastrophic Forgetting
How Neural ODEs Prevent Catastrophic Forgetting

Pavan Mohan
pavanmohan.netlify.app

December 23, 2024

What is a Traditional Neuron?

- A **traditional neuron** can be thought of as a **factory worker** performing specific tasks.
- Each **worker** (neuron) receives an **input**, processes it (calculations), and passes the result to the next worker.
- **Layers of neurons** represent different groups of workers. Each group performs its tasks and passes the results to the next.

How Does a Traditional Neuron Work?

- During **training**, the input data is fed into the first worker (input layer).
- Each worker applies transformations: multiplication, addition, and activation functions (e.g., ReLU, Sigmoid).
- After training, these transformations enable the network to make predictions.

Problem with Traditional Neurons

- Workers (neurons) are **not flexible** in their task handling.
- Once a worker learns one task, it's difficult for them to learn a new task without forgetting the previous one.
- This issue leads to **catastrophic forgetting**.

Catastrophic Forgetting Analogy

- Imagine a factory worker who is great at **assembling toys** (learning Task 1).
- The worker is asked to switch to **assembling cars** (learning Task 2).
- After learning how to assemble cars, the worker forgets how to assemble toys! They can only assemble cars now.

Why Does It Happen?

- Workers (neurons) in traditional networks cannot retain previous tasks while learning new ones.
- As new tasks are introduced, the model **overwrites** the previous knowledge.
- This makes sequential learning difficult because earlier tasks are lost.

- A **Neural ODE** neuron is like a **river** that flows continuously, rather than a factory with static workers.
- The river adapts its flow, changing course as it encounters new information, but it never completely forgets its past path.

How Neural ODE Neurons Work

- Neural ODEs operate as **continuous systems**—the model evolves over time.
- It adjusts to new data without **resetting** with each step, as traditional neurons do.
- The model **retains memory** while it learns new tasks.

Prevention of Catastrophic Forgetting

- Just like the river, Neural ODE neurons **do not forget** their past knowledge while adapting.
- New tasks are smoothly integrated without overwriting earlier tasks.
- The **path of the river** (past knowledge) is maintained as new information flows through.

Task Handling

- **Traditional Neuron**: Tasks are like steps in a checklist. New tasks overwrite old ones.
- **Neural ODE Neuron**: Tasks are like adjusting the flow of a river. New tasks adjust the flow, but previous tasks remain part of the journey.

Memory

- **Traditional Neuron**: Memory is like a hard drive—new information overwrites old data.
- **Neural ODE Neuron**: Memory is like a tree's growth rings—it evolves, but old information remains as part of the structure.

Learning Method

- **Traditional Neuron**: Learning happens in discrete steps, with each layer trained separately. New tasks overwrite previous tasks.
- **Neural ODE Neuron**: Learning happens through continuous updates, with the model constantly adjusting to new tasks without forgetting previous ones.

Real-World Example

- Imagine a robot learning two tasks:
    - Task 1: Picking up boxes.
    - Task 2: Stacking boxes.
- With traditional neurons, the robot forgets how to pick up boxes after learning to stack them.
- With Neural ODE neurons, the robot can **learn both tasks simultaneously**, retaining the ability to pick up boxes while stacking them.

- Traditional robot learning is like a **checklist**.
- New tasks overwrite previous ones.
- Neural ODE robot learning is like **continuous improvement**.
- Past knowledge is integrated without forgetting.

For more details: Refer to this GitHub Repo

Conclusion

- **Traditional Neurons**: Like discrete workers, they forget tasks when learning new ones.
- **Neural ODE Neurons**: Like a continuous river, they evolve with each task while retaining previous knowledge.
- Neural ODEs prevent **catastrophic forgetting** by adapting smoothly and integrating past knowledge.
- Neural ODEs are better suited for **sequential learning** tasks, maintaining performance across a series of tasks.