



# Chapter 2

## Evaluation of the Distributed Strategies for Data Parallel Deep Learning Model in TensorFlow

**Aswathy Ravikumar**

 <https://orcid.org/0000-0003-0897-6991>  
Vellore Institute of Technology, India

**Harini Sriraman**

 <https://orcid.org/0000-0002-2192-8153>  
Vellore Institute of Technology, India

### ABSTRACT

*Distributed deep learning is a branch of machine intelligence in which the runtime of deep learning models may be dramatically lowered by using several accelerators. Most of the past research reports the performance of the data parallelism technique of DDL. Nevertheless, additional parallelism solutions in DDL must be investigated, and their performance modeling for specific applications and application stacks must be reported. Such efforts may aid other researchers in making more informed judgments while creating a successful DDL algorithm. Distributed deep learning strategies are becoming increasingly popular as they allow for training complex models on large datasets in a much shorter time than traditional training methods. TensorFlow, a popular open-source framework for building and training machine learning models, provides several distributed training strategies. This chapter provides a detailed evaluation of the different TensorFlow strategies for medical data. The TensorFlow distribution strategy API is utilized to perform distributed training in TensorFlow.*

DOI: 10.4018/978-1-6684-9804-0.ch002

## **INTRODUCTION**

Deep learning has emerged as an effective method for dealing with a wide variety of challenging issues in a variety of fields, including machine vision, natural language processing, and recognition of speech, to name a few (Baby, 2014; Harini et al., 2022; John et al., 2021; Ravikumar et al., 2022; Robin et al., 2021). As deep learning models have increased in size and complexity over the past few years, it has become increasingly clear how essential it is to make use of distributed computing platforms to effectively train these models. According to Abadi et al. (2016), TensorFlow, considered one of the most renowned deep learning frameworks, provides customers with a distributed computing API. Thanks to this API, users can train large models in a distributed manner using several machines.

TensorFlow is widely employed in a wide variety of fields because to the numerous benefits it offers in deep learning (Muhammad Jaleed Khan et al., 2018; Zhang & Wei, 2020). This is because TensorFlow can be applied to a large range of problems. This is because Google was the one that initially built TensorFlow. On the other hand, TensorFlow is highly restricted for a single processing node, particularly when the size of the data set rises (Andrade & Trabasso, 2017; Bekeneva et al., 2020). As the size of the data set grows, the limits of TensorFlow become more and more apparent. How to overcome this barrier for TensorFlow to expand efficiently on ultra-large-scale systems, enhance the training pace of deep learning through parallelism, minimize the time required for training, improve the accuracy of both training and testing and apply deep learning to new domains. Sk et al. (2017) and Zhang et al. (2019) found that it significantly impacts the solving of challenging situations, which can have far-reaching implications.

Distributed data parallelism and distributed strategies are important concepts in machine learning when working with large datasets that cannot be processed on a single machine (Dean et al., n.d.). In TensorFlow, these concepts are implemented to enable training of machine learning models on distributed systems. Distributed data parallelism is a technique used to distribute the training of a model across multiple devices or machines. This is achieved by breaking the input data into multiple pieces and processing them in parallel on each device or machine. The model weights are then averaged across all devices or machines to produce the final model. Distributed strategies in TensorFlow are a set of tools and techniques used to distribute training and inference across multiple devices or machines. These strategies include data parallelism, model parallelism, and parameter servers. Data parallelism involves distributing the input data across multiple devices or machines and training the model on each device or machine. Model parallelism (Chen et al., 2019) involves distributing the model across multiple devices or machines and training different parts of the model on each device or machine. Parameter servers

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/evaluation-of-the-distributed-strategies-for-data-parallel-deep-learning-model-in-tensorflow/329545](http://www.igi-global.com/chapter/evaluation-of-the-distributed-strategies-for-data-parallel-deep-learning-model-in-tensorflow/329545)

## Related Content

---

### Comparative Analysis and Detection of Brain Tumor Using Fusion Technique of T1 and T2 Weighted MR Images

Padmanjali A. Hagargi (2021). *International Journal of Artificial Intelligence and Machine Learning* (pp. 54-61).

[www.irma-international.org/article/comparative-analysis-and-detection-of-brain-tumor-using-fusion-technique-of-t1-and-t2-weighted-mr-images/266496](http://www.irma-international.org/article/comparative-analysis-and-detection-of-brain-tumor-using-fusion-technique-of-t1-and-t2-weighted-mr-images/266496)

### Semantic Features Revealed in Consumer-Review Studies

Jing Yang, Tao Wuand Jeffrey Yi-Lin Forrest (2023). *Encyclopedia of Data Science and Machine Learning* (pp. 1143-1156).

[www.irma-international.org/chapter/semantic-features-revealed-in-consumer-review-studies/317515](http://www.irma-international.org/chapter/semantic-features-revealed-in-consumer-review-studies/317515)

### Forecasting Price of Amazon Spot Instances Using Machine Learning

Manas Malikand Nirbhay Bagmar (2021). *International Journal of Artificial Intelligence and Machine Learning* (pp. 71-82).

[www.irma-international.org/article/forecasting-price-of-amazon-spot-instances-using-machine-learning/277435](http://www.irma-international.org/article/forecasting-price-of-amazon-spot-instances-using-machine-learning/277435)

### Early Warning System Framework Proposal, Based on Big Data Environment

Goran Klepac, Robert Kopaland Leo Mrcic (2019). *International Journal of Artificial Intelligence and Machine Learning* (pp. 35-66).

[www.irma-international.org/article/early-warning-system-framework-proposal-based-on-big-data-environment/233889](http://www.irma-international.org/article/early-warning-system-framework-proposal-based-on-big-data-environment/233889)

### The Current Generation 5G and Evolution of 6G to 7G Technologies: The Future IoT

Sandeep Bhatia, Neha Goeland Soniya Verma (2023). *Handbook of Research on Machine Learning-Enabled IoT for Smart Applications Across Industries* (pp. 456-478).

[www.irma-international.org/chapter/the-current-generation-5g-and-evolution-of-6g-to-7g-technologies/326010](http://www.irma-international.org/chapter/the-current-generation-5g-and-evolution-of-6g-to-7g-technologies/326010)