

1 abstract

distributed online learning, each learner optimizes its own learning parameter based on local data source and correspond with neighbours timely.

privacy breaches(隐私泄露)

centralized approach

online data collection is inherently decentralized: data source are often widely distributed in different geographical locations.

wide distribution:online learning high velocity: high dimensionality: sparse, spark and parallelize privacy concern:

exchange intermediate parameters with a random part of their own neighboring(spark trees: find a leaf and route to the corresponding sub-tree)

distributed online learning

similar to the mini-batch online learning. each iteration, the process receives K instances. Then each node processes one of the instances and updates its local model. These nodes communicate with each other to keep the consistency of their local model.

A key factor in designing a distributed algorithm is the communication load between nodes.

2 INTRODUCTION

All nodes exchange their local parameter with their neighboring nodes.

For the privacy mechanism, differentially private framework is proposed with sensitive data